

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:19:50 ; Search time 49.13 Seconds
(Without alignments)

13.691 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28318 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	34.4	4	A40135	branched-chain-amino acid receptor be
2	11	34.4	5	PT0644	T-cell receptor be
3	11	34.4	5	B40817	34.5K structural p
4	11	34.4	5	D44817	35K structural p
5	10	31.2	7	E30608	Ig kappa chain V-I
6	9	28.1	3	T13892	Cytochrome-c oxida
7	9	28.1	5	E42364	flagellar protein
8	9	28.1	5	E60274	major protein anti
9	9	28.1	5	A37114	hypoxanthine phosph
10	9	28.1	5	H44817	34.5K structural p
11	9	28.1	5	F44817	34.5K structural p
12	9	28.1	5	S69237	surface protein te
13	9	28.1	6	A60986	N-formyl oligopept
14	9	28.1	6	S14159	paraspinal crystal
15	9	28.1	6	A43766	28K ubiquitin-immu
16	9	28.1	6	I37263	Y protein - human
17	9	28.1	6	I65546	MHC H2-L antigen -
18	9	28.1	7	S25266	pile protein - Esc
19	9	28.1	7	A25269	sex pheromone ccr1
20	9	28.1	7	A30812	sex pheromone ccr1
21	9	28.1	7	PT0649	alpha-dextrin endo
22	9	28.1	7	S09066	globulin IV alpha
23	9	28.1	7	PS0254	18K protein 5507 -
24	8	25.0	3	PT0622	T-cell receptor be
25	8	25.0	4	PT0696	T-cell receptor be
26	8	25.0	4	PT0645	T-cell receptor be
27	8	25.0	4	PT0712	T-cell receptor be
28	8	25.0	4	PT0698	T-cell receptor be
29	8	25.0	4	PT0551	T-cell receptor be

ALIGNMENTS

30	8	25.0	4	2	PT0697	T-cell receptor be
31	8	25.0	5	2	A60521	glycogen phosphory
32	8	25.0	5	2	I40698	biotin B - Citrobac
33	8	25.0	5	2	I39964	ribosomal protein
34	8	25.0	5	2	I39966	ribosomal protein
35	8	25.0	5	2	I39965	ribosomal protein
36	8	25.0	5	2	T10954	hypothetical prote
37	8	25.0	5	2	I50385	myosin light chain
38	8	25.0	5	2	PT0308	Ig heavy chain CRD
39	8	25.0	5	2	PT0596	T-cell receptor be
40	8	25.0	5	2	PT0610	T-cell receptor be
41	8	25.0	5	2	PT0597	T-cell receptor be
42	8	25.0	5	2	PT0600	T-cell receptor be
43	8	25.0	5	2	PT0729	T-cell receptor be
44	8	25.0	5	2	PT0624	T-cell receptor be
45	8	25.0	5	2	PT0686	T-cell receptor be

RESULT 1
A40135
branched-chain-amino-acid transaminase (EC 2.6.1.42), mitochondrial - rat (fragment)
N:Alternate names: branched-chain-amino-acid aminotransferase, mitochondrial
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-May-1992 #sequence_revision 13-May-1992 #text_change 30-Sep-1993
C:Accession: A40135
R:Hutson, S.M.; Wallin, R.; Hall, T.R.
submitted to the Protein Sequence Database, March 1992
A:Reference number: A40135
A:Accession: A40135
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <HUT>
C:Keywords: aminotransferase; mitochondrion

Query Match 34.4%; Score 11; DB 2; Length 4;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ISS 7
DB 1 VSS 3

RESULT 2

PT0644
T-cell receptor beta chain V-D-J region (111-16) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0644
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A:Reference number: PT0509; MIM: 91277601
A:Accession: PT0644
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEF>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 34.4%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FT 2
DB 4 FT 5

RESULT 3
B44817
34.5K structural protein - Leuconostoc oenos phage PZC11-15 (fragment)
C:Species: Leuconostoc oenos phage PZC11-15
C:Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C:Accession: B44817
R:Arendt, E.K.; Lonyaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.
A:Reference number: A44817; MUID:92085033
A:Accession: B44817
A:Molecule type: protein
A:Residues: 1-5 <AR6>
A:Note: sequence extracted from NCBI backbone (NCBIF:70342)

Query Match
Best Local Similarity 34.4%; Score 11; DB 2; Length 5;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 LTSS 7
1 11
Db 1 LATSS 5

RESULT 4
D44817
35K structural protein - Leuconostoc oenos phage P4C5-12 (fragment)
C:Species: Leuconostoc oenos phage P4C5-12
C:Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998
C:Accession: D44817
R:Arendt, E.K.; Lonyaud, A.; Hammes, W.P.
J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.
A:Reference number: A44817; MUID:92085033
A:Accession: D44817
A:Molecule type: protein
A:Residues: 1-5 <AR6>
A:Note: sequence extracted from NCBI backbone (NCBIF:70340)

Query Match
Best Local Similarity 34.4%; Score 11; DB 2; Length 5;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 LTSS 7
1 11
Db 1 LATSS 5

RESULT 5
E30608
19 kappa chain V-III region (Gag) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996
C:Accession: E30608
R:Goni, F.R.; Chen, P.P.; McGinits, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Sold
J. Immunol. 142, 3158-3163, 1989
A:Title: Structural and idiotype characterization of the L chains of human IgM autoanti
A:Reference number: A30601; MUID:89215279
A:Accession: E30608
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <GON>
A:Keywords: heterotrimer; immunoglobulin

Query Match
Best Local Similarity 31.2%; Score 10; DB 2; Length 7;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 LTSS 6

Db 4 LTOS 7

RESULT 6
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [Imported] - river lamprey mitochondrion (f
C:Species: mitochondrion lampetra fluviatilis (river lamprey)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: T13892
R:Delarbre, C.; Barriel, V.; Tiller, S.; Janyier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A:Title: The main features of the granulate mitochondrial DNA between the ND1 and the
A:Reference number: 217775; MUID:97398704
A:Accession: T13892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3
A:Cross-references: EMBL:Y09528; NID:92340016; PIDN:CAAT0721.1; PID:94379123
C:Genetics:
A:Genome: mitochondrion
A:Note: COI
C:Keywords: mitochondrion; oxidoreductase

Query Match
Best Local Similarity 28.1%; Score 9; DB 3; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
1 11
Db 2 TL 3

RESULT 7
E42364
flagellar protein fliR - Salmonella typhimurium (fragment)
C:Species: Salmonella typhimurium
C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C:Accession: E42364
R:Voelter, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A:Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and
A:Reference number: A42364; MUID:91258342
A:Accession: E42364
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5 <VOG>
A:Cross-references: GB:M62408

Query Match
Best Local Similarity 28.1%; Score 9; DB 2; Length 5;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
1 11
Db 3 TL 4

RESULT 8
E60274
major protein antigen MPT63 - Mycobacterium tuberculosis (fragment)
C:Species: Mycobacterium tuberculosis
C:Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C:Accession: E60274
R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A:Title: Isolation and partial characterization of major protein antigens in the cult
A:Reference number: A60274; MUID:91099899
A:Accession: E60274
A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-5 <NAG>

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTLT 4
: :
Db 2 YPIT 5

RESULT 9

A37114
hypoxanthine phosphoribosyltransferase (EC 2.4.2.8) - fluke (Schistosoma mansoni) (fragm

C:Species: Schistosoma mansoni
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 23-Jun-1993

C:Accession: A37114

R:Yuan, L.; Craig, S.P.; McKerrow, J.H.; Wang, C.C.

J. Biol. Chem. 265, 13528-13532, 1990
A:Title: The hypoxanthine-guanine phosphoribosyltransferase of Schistosoma mansoni. Fur

A:Reference number: A37114; MUID:90337955

A:Accession: A37114

A:Status: preliminary

A:Molecule type: protein
A:Residues: 1-5 <YUA>
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 ISS 7
: :
Db 1 MSS 3

RESULT 10

H44817
34.5K structural protein - Leuconostoc oenos phage P32 (fragment)

C:Species: Leuconostoc oenos phage P32

C:Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998

C:Accession: H44817

R:Arendt, E.K.; Lonnau, A.; Hammes, W.P.

J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.

A:Reference number: A44817; MUID:92085033

A:Accession: H44817
A:Molecule type: protein
A:Residues: 1-5 <ARE>
A>Note: sequence extracted from NCBI backbone (NCBIP:70332)

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LTISS 7
: :
Db 1 MATSS 5

RESULT 11

F44817
34.5K structural protein - Leuconostoc oenos phage P54 (fragment)

C:Species: Leuconostoc oenos phage P54

C:Date: 31-Mar-1993 #sequence_revision 22-May-1998 #text_change 22-May-1998

C:Accession: F44817

R:Arendt, E.K.; Lonnau, A.; Hammes, W.P.

J. Gen. Microbiol. 137, 2135-2139, 1991
A:Title: Lysogeny in Leuconostoc oenos.
A:Reference number: A44817; MUID:92085033

A:Accession: F44817
A:Molecule type: protein
A:Residues: 1-5 <ARE>
A>Note: sequence extracted from NCBI backbone (NCBIP:70335)

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 LTISS 7
: :
Db 1 MATSS 5

RESULT 12

S69237
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)

C:Species: Staphylothermus marinus

C:Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 17-Mar-1999

C:Accession: S69237

R:Peterson, J.; Nilsch, M.; Kuehlmoorgen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; Eng

J. Mol. Biol. 245, 385-401, 1995

A:Title: Tetrabrachion: a filamentous archaeobacterial surface protein assembly of unu

A:Reference number: S69237; MUID:95139068

A:Accession: S69237

A:Molecule type: protein

A:Residues: 1-5 <PET>

A:Experimental source: strain F1, DSM 3639

C:Keywords: cell wall; glycoprotein; heat-stable protein

Query Match 28.1%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
: :
Db 2 TL 3

RESULT 13
A60986
N-formyl oligopeptide - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C:Accession: A60986
R:Broom, M.F.; Mellor, D.M.; Chadwick, V.S.
Experientia 45, 1097-1099, 1989
A:Title: Purification and amino acid sequencing of naturally occurring N-formyl-methi

A:Reference number: A60986; MUID:90092408
A:Accession: A60986
A:Molecule type: protein
A:Residues: 1-6 <BRO>
A:Comment: This hexapeptide was the longest of several N-formyl oligopeptides reporte

F1/Modified site: N-formylmethionine #status experimental
Query Match 28.1%; Score 9; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
: :
Db 3 FTL 5

RESULT 14
S14159
parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)
N:Alternate names: delta-endotoxin
C:Species: Bacillus thuringiensis
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997

C:Accession: S14159
R:Convents, D.; Cherlet, M.; van Damme, J.; Lasters, I.; Lauwerijs, M.
Eur. J. Biochem. 195, 631-635, 1991
A:Title: Two structural domains as a general fold of the toxic fragment of the Bacillus
A:Reference number: S14087; MUID:91153300
A:Accession: S14159
A:Molecule type: protein
A:Residues: 1-6 <CON>

Query Match 28.1%; Score 9; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 ISS 7
II:
Db 1 IST 3

RESULT 15

A43766
28K ubiquitin-immunoreactive protein - inkx cap (Coprinus cinereus) (fragment)
C:Species: Coprinus cinereus
C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 30-Sep-1993
C:Accession: A43766
R:Kanda, T.; Inoue, M.; Akiyama, M.
Biochimie 72, 355-359, 1990
A:Title: Purification and characterization of an ubiquitin-immuno-reactive protein local
A:Reference number: A43766; MUID:91002724
A:Accession: A43766
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-6 <KAN>

Query Match 28.1%; Score 9; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
II:
Db 5 TL 6

Search completed: July 15, 2002, 13:27:07
Job time: 437 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:26:20 ; Search time 24.49 Seconds

(without alignments)
11.067 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTUTISS 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	37.5	7	1	GFRP_MOUSE
2	11	34.4	5	1	RE31_LITRU
3	11	34.4	5	1	RE32_LITRU
4	10	31.2	5	1	E104_LITRU
5	9	28.1	7	1	CCF1_ENTFA
6	9	28.1	7	1	CIA_ENTFA
7	8	25.0	5	1	ALL4_CARMA
8	8	25.0	5	1	BI09_CITFR
9	8	25.0	5	1	PSK_DAUCA
10	8	25.0	6	1	LOK1_LOCOMI
11	8	25.0	7	1	ALL2_CARMA
12	8	25.0	7	1	ALL3_CARMA
13	8	25.0	7	1	ALL4_CARMA
14	8	25.0	7	1	ALL5_CARMA
15	8	25.0	7	1	ALL7_CYPDO
16	8	25.0	7	1	UC24_MAIZE
17	8	25.0	4	1	EOS1_HUMAN
18	7	21.9	4	1	RM01_YEAST
19	7	21.9	5	1	BI0A_CITFR
20	7	21.9	6	1	UN06_CLOPA
21	7	21.9	4	1	ACH1_ACHFU
22	6	18.8	4	1	FAR3_HIRME
23	6	18.8	4	1	FAR4_HIRME
24	6	18.8	4	1	FEKA_AWTEL
25	6	18.8	4	1	FLRN_HIRME
26	6	18.8	4	1	FLRN_AWTEL
27	6	18.8	4	1	FMRE_MACNI
28	6	18.8	4	1	FYRI_AWTEL
29	6	18.8	4	1	OCPI_OCTMI
30	6	18.8	5	1	E103_LITRU
31	6	18.8	5	1	FARP_AWTR
32	6	18.8	5	1	PAP2_PARMA
33	6	18.8	5	1	RE11_LITRU

34	6	18.8	5	1	RE21_LITRU	P82071 litoria rub
35	6	18.8	5	1	SUGA_ACHDO	P19991 acheta dome
36	6	18.8	5	1	TPIS_CANFA	P54714 canis famli
37	6	18.8	5	1	UC22_MAIZE	P80628 zea mays (m
38	6	18.8	6	1	CIP1_MYTED	P13736 mytilus edu
39	6	18.8	6	1	CIP2_MYTED	P13737 mytilus edu
40	6	18.8	6	1	E101_LITRU	P82096 litoria rub
41	6	18.8	6	1	FARP_MONEX	P41866 monlexia ex
42	6	18.8	7	1	FAR1_ASCSU	P31889 ascaris suu
43	6	18.8	7	1	FAR1_HELTI	P41871 helisma tr
44	6	18.8	7	1	FAR1_PROCL	P38499 procambarus
45	6	18.8	7	1	FAR2_ASCSU	P31890 ascaris suu

ALIGNMENTS

```

RESULT 1
GFRP_MOUSE
ID GFRP_MOUSE STANDARD; PRT; 7 AA.
AC P99025;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN GCHFR OR GFRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Yan J.X.,
RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
RA Cowthorne M.;
RU Submitted (Aug-1998) to the SWISS-PROT data bank.
CC -!- FUNCTION: MEDIATES TETRAHYDROPTERIN INHIBITION OF GTP
CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLMALANINE
CC (BY SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC DR SWISS-2DPAGE; P99025; MOUSE.
FT INIT_MET 0
FT NON_TER 7
SQ SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;

```

Query Match 37.5%; Score 12; DB 1; Length 7;
Best local Similarity 60.0%; Pred. NO. 1e+05; 1; Indels 0; Gaps 0;
Matches 3; Conservative 1; Mismatches 1;

QY 3 LITSS 7
| | |
Db 3 LLIST 7

```

RESULT 2
RE31_LITRU
ID RE31_LITRU STANDARD; PRT; 5 AA.
AC P82072;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Robellidin 3.1.
DE Litorea rubella (Desert tree frog).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litorea.
OX NCBI_TaxId=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,

```

RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RL 'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
AU Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAH.
KW Amphibian skin; Amidation.
FT MOD.RES 5
SQ SEQUENCE 5 AA: 656 MW: 71A9C9CB10300000 CRC64:

Query Match 34.4%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
||
Db 4 FT 5

RESULT 3
RE32_LITRU STANDARD; PRT; 5 AA.
AC P82073;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidin 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:0-0(1999).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA: 570 MW: 71A9C9C862A00000 CRC64:

Query Match 34.4%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
||
Db 4 FT 5

RESULT 4
E104_LITRU STANDARD; PRT; 5 AA.
AC P82100;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Electrin 4.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;

RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:0-0(1999).
CC -1- SUBCELLULAR LOCATION: SECRETED.
KW Amphibian skin; Amidation.
FT MOD.RES 5
SQ SEQUENCE 5 AA: 616 MW: 61F2D1A059A00000 CRC64:

Query Match 31.2%; Score 10; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTI 5
:|:
Db 2 TTV 4

RESULT 5
CCF1_ENTFA STANDARD; PRT; 7 AA.
AC P20104;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CCF10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adelt J.C., Dunn G.M., Suzuki A.;
RT "Structure of CCF10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, pCF10.";
RL J. Biol. Chem. 263:14574-14578(1988).
CC -1- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC HEMOLYSIN PLASMID PCF10.
DR PIR, A30812; A30812.
KW Pheromone.
SQ SEQUENCE 7 AA: 790 MW: 72C9D2C731B2C740 CRC64:

Query Match 28.1%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
||
Db 3 TL 4

RESULT 6
CIA_ENTFA STANDARD; PRT; 7 AA.
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]

RP SEQUENCE.
 RX MEDLINE-87005252; PubMed-3092276;
 RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
 RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
 RT "Isolation and structure of the streptococcus faecalis sex pheromone,
 RT CAM373.";
 RL FEBS Lett. 206:69-72(1986).
 CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
 CC HARBORING PAM373.
 CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
 CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
 CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
 DR PIR: A25269; A25269.
 KW Pheromone.
 SQ SEQUENCE 7 AA; 734 MW; 758BD72059C05DB0 CRC64;

Query Match 28.1%; Score 9; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
 | |
 Db 3 FTL 5

RESULT 7
 AL14_CARMA STANDARD; PRT; 5 AA.
 AC P81817;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 14.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunidea; Portunidae; Carcinus.
 OX NCBI_TaxID-6759;
 RN [1];
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE-98121193; PubMed-9461295;
 RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation; Multigene family.
 KW MOD_RES 5 5 AMIDATION (POTENTIAL).
 FT MOD_RES 5 5
 SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 5;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
 | |
 Db 3 FTL 5

RESULT 8
 BIOB_CITFR STANDARD; PRT; 5 AA.
 AC P12997;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase) (Fragment).
 GN BIOB.

OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Citrobacter.
 OX NCBI_TaxID-546;
 RN [1];
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89006280; PubMed-2971595;
 RA Shtlun D., Campbell A.;
 RT "Transcriptional regulation and gene arrangement of *Escherichia coli*,
 RT Citrobacter freundii and *Salmonella typhimurium* biotin operons.";
 RL Gene 67:203-211(1988).
 CC -1- CATALYTIC ACTIVITY: Dehydrobiotin + sulfur = biotin.
 CC -1- PATHWAY: Biotin biosynthesis; last step.
 CC -1- SIMILARITY: BELONGS TO THE BIOTIN AND LIPOIC ACID SYNTHETASES
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M21922; -; NOT_ANNOTATED_CDS.
 DR Biotin biosynthesis; Iron-sulfur; Transferase.
 FT NON_TER 5
 FT SEQUENCE 5 AA; 532 MW; 75A5B1ED6F00000 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SS 7
 | |
 Db 4 SS 5

RESULT 9
 PSK_DAUCA STANDARD; PRT; 5 AA.
 AC PS8261;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phytosulfoline-alpha (PSK-alpha) [Contains: phytosulfoline-beta (PSK-
 DE beta)].
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID-4039;
 RN [1];
 RP SEQUENCE AND IDENTIFICATION BY MASS SPECTROMETRY.
 RC STRAIN=cv. US-Harumakigosun;
 RX MEDLINE-20212743; PubMed-10750705;
 RA Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
 RA Kanada H., Sakagami Y.;
 RT "A secreted peptide growth factor, phytosulfoline, acting as a
 RT stimulatory factor of carrot somatic embryo formation.";
 RL Plant Cell Physiol. 41:27-32(2000).
 CC -1- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
 CC CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
 CC EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
 CC PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOLINE FAMILY.
 CC Growth factor; Sulfation.
 KW PEPTIDE 1 4 PHYTOSULFOLINE-BETA.
 FT MOD_RES 1 1 SULFATION.
 FT MOD_RES 3 3 SULFATION.

SQ SEQUENCE 5 AA: 687 MW: 76C1B50AB300000 CRC64:

Query Match 25.0%; Score 8; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
1
3 YL 4

RESULT 10
LOR1_LOCM1 STANDARD: PRT: 6 AA.
ID LOR1_LOCM1
AC P41491;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Locustakinh I.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_Taxid=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=92262851; PubMed=1585017;
RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
de loof A.;
RT "Locustalin, a novel myotropic peptide from Locusta migratoria,
RT isolation, primary structure and synthesis.";
RL Regul. Pept. 37:49-57(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION
CC OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN
CC TUBULES.
CC PIR: A61068; A61068.
DR Neuropeptide: Amidation.
KW MOD_RES 6
FT SEQUENCE 6 AA: 654 MW: 686365MSB9CDB000 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SS 7
1
3 SS 4

RESULT 11
ALL2_CARMA STANDARD: PRT: 7 AA.
ID ALL2_CARMA
AC P81805;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 2.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_Taxid=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorppe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";

RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7
SQ SEQUENCE 7 AA: 770 MW: 672879CDBC85DB70 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
1
5 FGL 7

RESULT 12
ALL3_CARMA STANDARD: PRT: 7 AA.
ID ALL3_CARMA
AC P81806;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 3.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_Taxid=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorppe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA: 796 MW: 672879CDBC8476B70 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
1
5 FGL 7

RESULT 13
ALL4_CARMA STANDARD: PRT: 7 AA.
ID ALL4_CARMA
AC P81807;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 4.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_Taxid=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorppe A.;

RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab *Carcinus maenas*.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
KW -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Multigene family.
SQ SEQUENCE 7 AA: 782 MW: 672879CDBC476AC0 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
| |
| |
Db 5 FGL 7

RESULT 14
ALL5_CARMA STANDARD; PRT; 7 AA.
AC P81808: 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab *Carcinus maenas*.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
KW -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Amidation; Multigene family.
FT MOD_RES 7
SQ SEQUENCE 7 AA: 781 MW: 672879CDBC476420 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
| |
| |
Db 5 FGL 7

RESULT 15
ALL7_CYPDPO STANDARD; PRT; 7 AA.
AC P82158: 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Codling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;

RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA: 873 MW: 672879CDB569350 CRC64;

Query Match 25.0%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
| |
| |
Db 5 FGL 7

Search completed: July 15, 2002, 13:38:10
Job time: 710 sec

•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:25:30 ; Search time 86.73 Seconds
(without alignments)
13.962 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTLRIS 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 65

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phase:*
11: sp_plant:*
12: sp_ricent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_rviro:*
17: sp_bacteriap:*
18: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	50.0	7	2	P70804
2	12	37.5	7	2	O34028
3	11	34.4	5	13	P82072
4	11	34.4	5	13	P82073
5	10	31.2	7	2	O54248
6	10	31.2	5	13	P82100
7	9	28.1	7	2	O07354
8	9	28.1	7	10	P82445
9	8	25.0	6	10	P82181
10	8	25.0	6	10	P82182
11	8	25.0	7	2	P72081
12	7	21.9	7	2	O47029
13	7	21.9	7	12	O65578
14	6	18.8	5	13	P82070
15	6	18.8	5	13	P82071
16	6	18.8	5	13	P82099

17	6	18.8	6	10	P82541	P82541 spinacia ol
18	6	18.8	6	13	P82096	P82096 litoria rub
19	6	18.8	7	2	O50556	O50556 actinobacil
20	6	18.8	7	4	O15903	O15903 homo sapien
21	6	18.8	7	8	O95945	O95945 saccharomyc
22	6	18.8	7	8	O98666	O98666 spinacia ol
23	6	18.8	7	10	O9c5B3	O9c5B3 arabidopsis
24	6	18.8	7	12	O9YQ10	O9YQ10 transmissib
25	6	18.8	7	13	O42564	O42564 ligu rubrip
26	6	18.8	7	13	P82065	P82065 litoria rub
27	6	18.8	5	2	P83073	P83073 bacillus ce
28	5	15.6	7	10	O47505	O47505 escherichia
29	5	15.6	7	11	O63668	O63668 rattus norv
30	5	15.6	7	12	O67113	O67113 influenza a
31	5	15.6	7	13	P82101	P82101 litoria rub
32	5	15.6	7	13	O07624	O07624 rous sarcom
33	5	15.6	7	11	O08433	O08433 rattus norv
34	4	12.5	7	4	O15897	O15897 homo sapien
35	4	12.5	7	8	P92421	P92421 psathyrosta
36	4	12.5	7	8	P92385	P92385 hordeum mar
37	4	12.5	7	8	P92372	P92372 haynaldia v
38	4	12.5	7	8	P92403	P92403 lophopyrum
39	4	12.5	7	8	P92425	P92425 pseudorogn
40	4	12.5	7	8	P92387	P92387 pentadactyl
41	4	12.5	7	8	P92427	P92427 peridictyon
42	4	12.5	7	8	P92390	P92390 heteranthel
43	4	12.5	7	8	P92226	P92226 crithopsis
44	4	12.5	7	8	P92214	P92214 amblyopyrum
45	4	12.5	7	8	P92214	P92214 amblyopyrum

ALIGNMENTS

RESULT 1
ID P70804 PRELIMINARY; PRT; 7 AA.
AC P70804: 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALGT PROTEIN (FRAGMENT).
GN ALGT.
OS Acetobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Acetobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Rehm B.H.A., Ertesvag H., Vaila S.;
RT "A new Acetobacter vinelandii mannuronan C-5-epimerase gene (algc) is
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL, X67973; CAA61230.1; -;
FT NON_TER 1
SQ SEQUENCE 7 AA: 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 50.0%; Score 16; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 4 TISS 7
1:11
Db 2 TVSS 5
RESULT 2
O34028 PRELIMINARY; PRT; 7 AA.

AC 034026;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CAECHOL-2,3-DIOXYGENASE (FRAGMENT).
GN PINE.
OS Sphingomonas chungbukensis.
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
OC Sphingomonas.
OX NCBI_TaxID=56193;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DU77;
RA Kim Y.-C.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88298; AAB6311.1; -.
KW Dioxigenase.
FT NON-TER
SQ SEQUENCE 7 AA; 868 MW; 71AA52D1A699D460 CRC64;

Query Match 37.5%; Score 12; DB 2; Length 7;
Best Local Similarity 20.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTIS 7
Db 3 MTVNT 7

RESULT 3
P82072 PRELIMINARY; PRT; 5 AA.

AC P82072;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN SECRETION;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowle J.H., Gao C.,
Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
RT Aust. J. Chem. 49:955-963(1996).
CC -I- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -I- MASS SPECTROMETRY: MW=655; METHOD=FAE.
KW Amphibian skin; Amidation.
FT MOD_RES 5
SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 34.4%; Score 11; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
Db 4 FT 5

RESULT 4
P82073 PRELIMINARY; PRT; 5 AA.

AC P82073;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RUBELLIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN SECRETION;
RA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electrica. Comparison with the skin peptides from Litoria
rubella.";
RT Aust. J. Chem. 52:0-0(1999).
CC -I- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9CB862A000000 CRC64;

Query Match 34.4%; Score 11; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
Db 4 FT 5

RESULT 5
Q54248 PRELIMINARY; PRT; 7 AA.

AC Q54248;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RPIO PROTEIN (FRAGMENT).
GN RPIO.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2-3-11;
RX MEDLINE=20011291; PubMed=10542330;
RA Poehling S., Plepersberg W., Wehmeler U.F.;
RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
N2-3-11 and interaction of the sec Y protein with the SecA protein.";
RL Biochim. Biophys. Acta 1447:298-302(1999).
DR EMBL; X95915; CAA65160.1; -.
FT NON-TER
SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 34.4%; Score 11; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TLT 4
Db 2 TVT 4

RESULT 6
P82100 PRELIMINARY; PRT; 5 AA.

ID P82100;
AC P82100;

DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hyllidae;
 OC Litoria.
 NC NCB1_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN SECRETION;
 RA Maniatis P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 Litoria electrica. Comparison with the skin peptides from Litoria
 rubella".
 RL Aust. J. Chem. 52:0-0(1999).
 KW Amphibian skin; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 31.2%; Score 10; DB 13; Length 5;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTI 5
 :1:
 DB 2 ITV 4

RESULT 7
 ID 007354 PRELIMINARY; PRT; 7 AA.
 AC 007354
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE NTRK (FRAGMENT).
 GN NTRK.
 OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanotheca PCC 8801).
 OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
 NC NCB1_TaxID=41431;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RF-1;
 RX MEDLINE=99231861; PubMed=10217509;
 RA Huang T.C., Lin K.F., Chu M.K., Chen H.M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RF-1".
 RL Microbiology 145:743-753(1999).
 DR EMBL: AF003700; AAC35193.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match 28.1%; Score 9; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
 :1:
 DB 3 FDL 5

RESULT 8
 ID P82445 PRELIMINARY; PRT; 7 AA.
 AC P82445
 DT 01-JUN-2000 (TREMBlrel. 14, Created)
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE 10 KDA CELL WALL PROTEIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; easterids I; Solanales; Solanaceae; Nicotiana.
 NC NCB1_TaxID=4097;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. PETT HAVANA;
 RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
 RA Wojtaszek P., Bolwell G.P.;
 RT "Proteomic study of secondary cell wall proteins from transformed
 tobacco culture".
 RL Planta 0:0-0(2000).
 CC -1- SUBCELLULAR LOCATION: CELL WALL.
 CC -1- TISSUE SPECIFICITY: XYLEM.
 KW Cell wall.
 FT NON_TER 7
 SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;

Query Match 28.1%; Score 9; DB 10; Length 7;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTI 5
 :1:
 DB 1 VTY 3

RESULT 9
 ID P82181 PRELIMINARY; PRT; 6 AA.
 AC P82181
 DT 01-JUN-2000 (TREMBlrel. 14, Created)
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiales; Spinacia.
 NC NCB1_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. ALVARO; TISSUE=LEAF;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Sudramanatan A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 the 50 S subunit of an organelle ribosome (chloroplast)".
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR001790; Ribosomal_L10.
 DR InterPro: IPR002363; Ribosomal_L10_eub.
 DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
 DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 25.0%; Score 8; DB 10; Length 6;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 IS 6
 :1:
 DB 2 IS 3

```
RESULT 10
ID P82182 PRELIMINARY; PRT; 6 AA.
AC P82182;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 GAMMA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
OC Caryophyllales: Caryophyllales: Chenopodiaceae: Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALMARO; TISSUE=LEAF;
RX MEDLINE=20435798; Pubmed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 273:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; RIBOSOMAL_L10.
DR InterPro: IPR002363; RIBOSOMAL_L10_eub.
DR Pfam: PF00466; RIBOSOMAL_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; chloroplast; rRNA-binding.
FT NON_TER
SQ SEQUENCE 6 AA: 675 MW: 6321B435B05DB000 CRC64;

Query Match 25.0%; Score 8; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 IS 6
   11
Db 2 IS 3

RESULT 11
ID P72081 PRELIMINARY; PRT; 7 AA.
AC P72081;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE 3'-METHYLCEPHEM HYDROXYLASE (FRAGMENT).
OS CEPF.
GN Nocardia lactamdurans.
OC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
OC Actinomycetales: Pseudonocardiaceae; Pseudonocardiaceae; Amycolatopsis.
OX NCBI_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009872; Pubmed=7557411;
RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA Liras P.;
RT "Characterization of the cmcH genes of Nocardia lactamdurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL: Z21682; CAA79797.1; -.
FT NON_TER
SQ SEQUENCE 7 AA: 746 MW: 75A1B2CDD1E681C0 CRC64;

Query Match 25.0%; Score 8; DB 2; Length 7;
Matches 2; Conservative 0; Mismatches 7;
```

```
Best Local Similarity 33.3%; Pred. No. 5.6e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 ISS 7
   11
Db 5 VTS 7

RESULT 12
ID Q47029 PRELIMINARY; PRT; 7 AA.
AC Q47029;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, last annotation update)
DE AAD A1 PROTEIN (FRAGMENT).
GN AAD A1.
OS Enterobacter cloacae.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
OC Enterobacter.
OX NCBI_TaxID=550;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94079349; Pubmed=8257126;
RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
RT "Analysis of the aac(3)-Via gene encoding a novel 3-N-
RT acetyltransferase.";
RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
DR EMBL: M88012; AAA16193.1; -.
FT NON_TER
SQ SEQUENCE 7 AA: 744 MW: 633862D2C321A030 CRC64;

Query Match 21.9%; Score 7; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LT 4
   11
Db 1 IT 2

RESULT 13
ID Q65578 PRELIMINARY; PRT; 7 AA.
AC Q65578;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHEICAL 0.8 KDA PROTEIN (FRAGMENT).
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10320;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=COOPER;
RC MEDLINE=95313343; Pubmed=7793062;
RA Vicek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyzer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
DR EMBL: Z48053; CAA88130.1; -.
FT NON_TER
SQ SEQUENCE 7 AA: 758 MW: 6DD3345B1F1B1C0 CRC64;

Query Match 21.9%; Score 7; DB 12; Length 7;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 4 TISS 7
1
DB 1 TTHS 4

RESULT 14

P82070 PRELIMINARY; PRT; 5 AA.
AC P82070;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE RUBELLIDIN 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION.
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -!- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

Query Match 18.8%; Score 6; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 F 1
1
DB 3 F 3

RESULT 15

P82071 PRELIMINARY; PRT; 5 AA.
AC P82071.
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE RUBELLIDIN 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION.
RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

Query Match 18.8%; Score 6; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 F 1
1
DB 3 F 3

Search completed: July 15, 2002, 13:37:39
Job time: 729 sec

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands SI-S6, and may also include beta strands AI-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets SI/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands AI-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC AI/SI/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

Sequence 7 AA:

Query Match 87.5%; Score 28; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLLTIS 6
| | | | |
Db 2 flltis 7

RESULT 2

AAB30076
ID AAB30076 standard; Peptide: 7 AA.

AC AAB30076;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S4 peptide SEQ ID NO: 137.

KM Human: CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KM SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KM diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO200060070-A1.

PD 12-OCT-2000.

PE 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
* fragments

PS Disclosure: Page 15; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.

Sequence 7 AA:

Query Match 87.5%; Score 28; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLLTIS 6
| | | | |
Db 2 flltis 7

RESULT 3

AAY40737
ID AAY40737 standard; peptide; 7 AA.

AC AAY40737;

DT 01-DEC-1999 (first entry)

DE S4 derivative #11, beta strand of scaffold protein structure.

KM Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KM tumour; chemotherapeutic agent.

OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PE 31-MAR-1998; 98BP-0870065.

PR 31-MAR-1998; 98BP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines
*
PS Disclosure: Page 6; 105pp; English.

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands SI-S6, and may also include beta strands AI-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets SI/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands AI-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC AI/SI/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will

target the complex to tumour cells. Another surface may be bound to a cytotoxic molecule or an autoimmune antibody which may then kill the tumour cells. Therefore the scaffold protein may be used to target chemotherapeutic agents to specific cells. It may also be used to stabilize individual peptides in a peptide library and may be used in diagnostic techniques, and to stabilize antigens used as vaccines.

Sequence 7 AA:

Query Match 78.1%; Score 25; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTIS 6
:|||||
DB 2 ytlfts 7

RESULT 4

AAB30075
ID AAB30075 standard; Peptide; 7 AA.

AC AAB30075;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S4 peptide SEQ ID NO: 136.

KM Human; C11A-4; scaffold protein; antigen-binding; receptor-binding;
KM SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KM diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO20060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI; 2000-665002/64.

PT Scaffold composed of single-chain polypeptide having beta sandwich architecture carrying new and randomized peptide sequences useful as PT supporting framework and carrying antigen- or receptor binding fragments.

PS Disclosure; Page 15; 68pp; English.

CC The present invention is concerned with producing scaffold proteins CC based upon the human C11A-4 SCA domain. These scaffold proteins can be CC used as a scaffold to bind antigen- or receptor-binding fragments. These CC can be used in the treatment of diseases such as cancer, atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the CC production of the proteins of the invention.

Sequence 7 AA:

Query Match 78.1%; Score 25; DB 21; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTIS 6
:|||||

DB 2 ytlfts 7

RESULT 5

AAY40736
ID AAY40736 standard; peptide; 7 AA.

AC AAY40736;

DT 01-DEC-1999 (first entry)

DE S4 derivative #10, beta strand of scaffold protein structure.

KM Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KM tumour; chemotherapeutic agent.

OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI; 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure; Page 6; 105pp; English.

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a CC beta strand peptide which forms part of a beta sheet. Peptides CC (AAY40601-Y40605) together form a single-chain scaffold protein which CC contains at least 1 disulfide bond, contains less than 10% alpha helix CC and contains at least 6 beta-strands. The scaffold protein is constructed CC of beta strands S1-S6, and may also include beta strands A1-A3, or any CC functionally equivalent derivative of these sequences. The beta strands CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to CC the next by hydrogen bonds, which generate a beta sandwich architecture. CC If the additional beta strands A1-A3 are included in the structure the CC scaffold is constructed of two beta sheets, with the structures CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each CC other via amino acid loops, where at least one of the loops binds to a CC receptor or antigen. The scaffold protein is used to stabilize antigens CC or whole proteins such as receptors, or their fragments. It may be used CC to bind two separate molecules. For example, one surface of the scaffold CC may be bound to a protein which binds to a tumour antigen. This will CC target the complex to tumour cells. Another surface may be bound to a CC cytotoxic molecule or an autoimmune antibody which may then kill the CC tumour cells. Therefore the scaffold protein may be used to target CC chemotherapeutic agents to specific cells. It may also be used to CC stabilize individual peptides in a peptide library and may be used in CC diagnostic techniques, and to stabilize antigens used as vaccines.

Sequence 7 AA:

Query Match 75.0%; Score 24; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTIS 6
:|||||
DB 2 ftlfts 7

RESULT 6

AAB30074

ID AAB30074 standard; Peptide: 7 AA.

AC AAB30074;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S4 peptide SEQ ID NO: 135.

Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO200060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

PT Scaffold composed of single-chain polypeptide having beta sandwich
architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
fragments.

PS Disclosure: Page 15; 68pp; English.

CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCH domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-A29939 were used in the
CC production of the proteins of the invention.

SQ Sequence 7 AA;

Query Match 75.0%; Score 24; DB 21; Length 7;

Best Local Similarity 83.3%; Pred. No. 6.4e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1;

QY 1 FTLTIS 6

Db 2 ftlstis 7

RESULT 7

AAV40735

ID AAV40735 standard; Peptide: 7 AA.

AC AAV40735;

DT 01-DEC-1999 (first entry)

DE S4 derivative #9, beta strand of scaffold protein structure.

Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO200060070-A1.

PN EP94/582-A1.

XX 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure: Page 6; 105pp; English.

CC Sequences AAV40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAV40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAV40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

SQ Sequence 7 AA;

Query Match 65.6%; Score 21; DB 20; Length 7;

Best Local Similarity 66.7%; Pred. No. 6.4e+05; Mismatches 0; Indels 0; Gaps 0;

Matches 4; Conservative 2;

QY 1 FTLTIS 6

Db 2 ftlstis 7

RESULT 8

AAB30073

ID AAB30073 standard; Peptide: 7 AA.

AC AAB30073;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S4 peptide SEQ ID NO: 134.

Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO200060070-A1.

.
 XX 12-OCT-2000.
 PD
 XX
 PF 01-APR-1999; 99WO-EP02283.
 XX
 PR 01-APR-1999; 99WO-EP02283.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
 XX
 DR WPI; 2000-665002/64.
 XX
 PT Scaffold composed of single-chain polypeptide having beta sandwich
 PT architecture carrying new and randomized peptide sequences useful as
 PT supporting framework and carrying antigen- or receptor binding
 PT fragments -
 XX
 PS Disclosure: Page 15; 68pp; English.
 XX
 CC The present invention is concerned with producing scaffold proteins
 CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
 CC used as a scaffold to bind antigen- or receptor-binding fragments. These
 CC can be used in the treatment of diseases such as cancer,
 CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
 CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
 CC production of the proteins of the invention.
 XX
 SO Sequence 7 AA;

 Query Match 65.6%; Score 21; DB 21; Length 7;
 Best Local Similarity 66.7%; Pred. No. 6,4e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 Oy 1 FTLTIS 6
 Db 2 ystltis 7

 RESULT 9
 AAM51422
 ID AAM51422 standard; peptide; 6 AA.
 XX
 AC AAM51422;
 XX
 DT 08-JAN-2002 (first entry)
 XX
 DE Integrin activating peptide SEQ ID NO 1.
 XX
 KW Integrin stimulant; vulnery; injury healing;
 KM postsurgical tissue recovery.
 XX
 OS Unidentified.
 XX
 PN JP2001213898-A.
 XX
 PD 07-AUG-2001.
 XX
 PF 31-JAN-2000; 2000JP-0022469.
 XX
 PR 31-JAN-2000; 2000JP-0022469.
 XX
 PA (HISM) HISAMITSU PHARM CO LTD.
 XX
 DR WPI; 2001-629610/73.
 XX
 PT An injury healing and postsurgical tissue recovering integrin
 PT activating peptide -
 XX
 PS Claim 1; Page 3; 11pp; Japanese.
 XX
 CC The invention relates to novel peptides with vulnery activity, useful

CC for injury healing and postsurgical tissue recovery by acting as an
 CC integrin stimulant.
 CC
 XX
 SO Sequence 6 AA;

 Query Match 59.4%; Score 19; DB 22; Length 6;
 Best Local Similarity 60.0%; Pred. No. 6,4e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

 Oy 1 FTLTIF 5
 Db 1 yltlti 5

 RESULT 10
 AAM58711
 ID AAM58711 standard; peptide; 7 AA.
 XX
 AC AAM58711;
 XX
 DT 17-SEP-1998 (first entry)
 XX
 DE Tryptic 40 kD subunit CLMF peptide off PVDF #4.
 XX
 DE Cytotoxic lymphocyte maturation factor; CLMF; cytokine; synergize;
 KM Interleukin 2; IL-2; human; lymphoblastoid cell; detection; purifying;
 KM proliferation; cytotoxic T cell; transplantation; antibody.
 XX
 OS Homo sapiens.
 XX
 PN US5780597-A.
 XX
 PD 14-JUL-1998.
 XX
 PE 02-JUN-1995; 95US-0460061.
 XX
 PR 24-MAR-1992; 92US-0857023.
 PR 22-DEC-1989; 89US-0455708.
 PR 09-MAY-1990; 90US-0520935.
 PR 27-AUG-1990; 90US-0572284.
 PR 02-MAR-1994; 94US-0205011.
 PR 02-JUN-1995; 95US-0460061.
 XX
 PA (HOFF) HOFFMANN LA ROCHE INC.
 XX
 PI Chizzone RA, Gately MK, Gubler UA, Hulmes JD;
 PI Pan YE, Podlaski FJ, Stern AS;
 XX
 DR WPI; 1998-413150/35.
 XX
 PT New antibodies to cytotoxic lymphocyte maturation factor - useful
 PT for detecting, purifying, and/or blocking proliferation and
 PT activation of cytotoxic T cells, such as in transplantation(s)
 XX
 PS Example 3; Column 23; 71pp; English.
 XX
 CC An isolated antibody has been developed which binds specifically to
 CC cytotoxic lymphocyte maturation factor (CLMF) and neutralises CLMF
 CC induced proliferation by more than 50% in a CLMF dependent T cell
 CC growth assay and/or inhibits binding of at least 60% of the factor
 CC to phytohemagglutinin (PHA) activated peripheral blood lymphocyte
 CC (PBL) blasts as determined in a CLMF receptor binding assay. CLMF is
 CC a heterodimeric protein having a molecular weight band of 75 kD,
 CC determined by sodium dodecyl sulphate-polyacrylamide gel electrophoresis
 CC (SDS-PAGE) under non-reducing and/or reducing conditions, providing a
 CC first subunit having a molecular weight band of 40 kD and a second
 CC subunit having a molecular weight of 35 kD. The present sequence
 CC represents a tryptic 40 kD CLMF peptide off polyvinylidene difluoride
 CC (PVDF), from an example of the present invention. The antibody can be
 CC used for the purification and/or detection of CLMF. It is also used in
 CC therapeutic treatments which require selective blocking of proliferation
 CC and activation of cytotoxic T cells (CTLs) such as in transplantation.

XX Sequence 7 AA:
SQ
Query Match 56.2%; Score 18; DB 19; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 TLTI 5
1 1 1 1
DB 1 tltl 4

RESULT 11
AAV64225
AAV64225 standard; Peptide: 7 AA.
AC AAV64225:
DT 02-MAR-2000 (first entry)
DE Cadherin-related neuronal receptor CAR sequence SEQ ID NO:637.
DE
KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW Inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; protocadherin;
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; Obesity;
KW Rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurological disease.
XX
XX
OS Synthetic.
OS Homo sapiens.
PN MO9957149-A2.
PD 11-NOV-1999.
PF 05-MAY-1999; 99MO-CA00363.
PR 05-MAY-1998; 98US-0073040.
PR 06-NOV-1998; 98US-0187859.
PR 20-JAN-1999; 99US-0234395.
PR 08-MAR-1999; 99US-0264516.
PA (ADHE-) ADHEREX TECHNOLOGIES INC.
PI Blaschuk OM, Gour BJ, Byers S;
DR WPI: 2000-038791/03.
XX
XX
PT New cadherin modulating agents, used for modulating nonclassical
PT cadherin-mediated functions for treating e.g. cancers, obesity,
PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
PT disease -
XX
XX
PS Claim 99; Page 216; 252pp: English.

The present invention describes cadherin modulating agents (MA)
comprising peptides which comprise a nonclassical cadherin cell adhesion
recognition (CAR) sequence. The MAs can be used for modulating
nonclassical cadherin-mediated functions. They can be used for e.g.
inhibiting adhesion of nonclassical-cadherin expressing cells in a
mammal, enhancing delivery of a drug through the skin of a mammal,
enhancing delivery of a drug to a tumour in a mammal, treating cancer in
a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
expressing cell, preventing or treating obesity in a mammal, stimulating
blood vessel regression in a mammal, enhancing drug delivery to the
central nervous system, treating a demyelinating neurological disease,
increasing vasopermeability in a mammal, enhancing adhesion of
nonclassical cadherin-expressing cells, inhibiting synaptic stability in

CC a mammal, or preventing pregnancy in a mammal. They can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound
CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
CC in a mammal. They can also be used for treating e.g. psoriasis,
CC arthritis, age-related macular degeneration, multiple sclerosis and
CC diabetes. The products can also be used for detection and diagnosis and
CC in bioreactors. AAV60592 to AAV64572 represent specifically claimed
CC peptides, and AAV64573 to AAV64643 and AA233183 to AA233186 represent
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 7 AA:
Query Match 56.2%; Score 18; DB 21; Length 7;
Best Local Similarity 57.1%; Pred. No. 6.4e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 FTLTSS 7
1 1 1 1 1
DB 1 tltlss 7

RESULT 12
AAW39453
AAW39453 standard; peptide: 6 AA.
AC AAW39453:
DT 11-JUN-1998 (first entry)
DE Human T cell epitope proteolytic cleavage site 4.
DE
KW T cell epitope; immune response; human leukocyte antigen; HLA Class I;
KW vaccine; immunogenic; major histocompatibility complex; MHC; B cell;
KW disease; anti-tumour; anti-viral.
XX
XX
OS Synthetic.
OS Homo sapiens.
PN WO9741440-A1.
PD 06-NOV-1997.
PF 28-APR-1997; 97MO-NL00229.
PR 23-DEC-1996; 96EP-0203670.
PR 26-APR-1996; 96EP-0201145.
PA (UYLE-) RIJKSUNIV LEIDEN.
PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
PI Kast WM, Melief CJM, Offringa R, Toes REM, Van Der Burg SH;
DR WPI: 1997-549891/50.
XX
XX
PT Method of selecting T cell peptide epitope(s) - by measuring the
PT stability of HLA class I-peptide complexes on intact B cells
PT
XX
XX
PS Example 5; Page 42; 109pp: English.

Peptides AAW39450-W39734 are used in a novel method for the selection of
immunogenic T-cell peptide epitopes present in polypeptide antigens. The
method involves the identification of peptide sequences capable of
binding to an HLA (human leukocyte antigen) class I molecule and
measuring the binding of this epitope peptide to the HLA class I
peptide. The stability of binding of the peptide and MHC (major
histocompatibility complex) class I molecule is measured on intact human
B cells carrying the MHC molecule at their cell surfaces. The method can
be used to select peptide epitopes for generating vaccines against a
disease associated with the polypeptide, e.g. cancers or AIDS. The
peptide epitopes are especially T-cell peptide epitopes with strong
anti-tumour and anti-viral immune responses. Peptides AAW39450-W39503 are
proteolytic cleavage sites capable of linking together T cell epitopes in

CC vaccines.
XX
SQ Sequence 6 AA;
Query Match 53.1%; Score 17; DB 18; Length 6;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 FTLTIS 6
1 1 1 1 1 1
Db 1 ftlis 6
RESULT 13
AAW75358
ID AAW75358 standard; peptide; 6 AA.
XX
AC AAW75358;
XX
DT 02-FEB-1999 (first entry)
XX
DE Hexapeptide #13 binds anti-Ad5 fibre head MAb 7A2.7.
XX
DE Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;
XX
KW viral surface protein; hexapeptide expression library; adenovirus;
KW major histocompatibility complex; MHC; fibronectin; gene therapy;
KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.
XX
OS Synthetic.
XX
FN FR2758821-A1.
XX
PD 31-JUL-1998.
XX
PF 30-JAN-1997; 97FR-0001005.
XX
PR 30-JAN-1997; 97FR-0001005.
XX
PR 09-SEP-1997; 97FR-0011166.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Boulanger P, Hong SS, Karayan L;
XX
DR WPI; 1998-416458/36.
XX
PT Use of peptide(s) for facilitating or modulating attachment of
PT adenovirus to cells - useful for, e.g. treating or preventing
PT infection and improving uptake of gene therapy vectors
XX
PS Example 2; Fig 1; 48pp; French.
XX
CC The invention relates to methods for selecting and identifying a
CC cellular receptor for a virus, by immobilising, on a support, a
CC monoclonal antibody targeted to a viral surface protein that determines
CC attachment of the virus to the receptor. The immobilised antibody is the
CC incubated with a hexapeptide expression library and peptides bound to the
CC immobilised antibody are eluted by competitive binding with recombinant
CC fragments of the viral surface protein. In a reverse method, the viral
CC surface protein is immobilised and incubated with the peptide library. In
CC this case, the bound peptides are eluted by competitive binding with the
CC monoclonal antibody. The hexapeptides AAW75346-W75360 represent
CC peptides isolated by biopanning the library with the immobilised
CC monoclonal antibody (MAb) 7A2.7 and eluted using an adenovirus type 5
CC fibre head protein (AAW75335).
CC The methods are used to identify peptides from MHC Class I and/or
CC type III fibronectin proteins that allow or facilitate attachment by
CC adenovirus (Ad) to host cells and/or entry into the cells, and to
CC identify ligands that modulate Ad infection mediated by these peptides,
CC e.g. to treat or prevent Ad infections or to facilitate infection by Ad
CC gene therapy vectors used to treat genetic diseases, acquired immune
CC deficiency syndrome or cancer.

SQ Sequence 6 AA;
Query Match 53.1%; Score 17; DB 19; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FTLT 4
1 1 1 1 1 1
Db 3 yltt 6
RESULT 14
AAW75290
ID AAW75290 standard; peptide; 6 AA.
XX
AC AAW75290;
XX
DT 02-FEB-1999 (first entry)
XX
DE Hexapeptide #13 binds anti-Ad5 fibre head MAb 7A2.7.
XX
DE Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;
XX
KW viral surface protein; hexapeptide expression library; adenovirus;
KW major histocompatibility complex; MHC; fibronectin; gene therapy;
KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.
XX
OS Synthetic.
XX
FN FR2758885-A1.
XX
PD 31-JUL-1998.
XX
PF 30-JAN-1997; 97FR-0001005.
XX
PR 30-JAN-1997; 97FR-0001005.
XX
PR 10-JUL-1997; 97FR-0008796.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Boulanger P, Hong SS, Karayan L;
XX
DR WPI; 1998-416493/36.
XX
PT Selection and identification of cellular receptors for viruses -
PT used to control attachment and entry of adenovirus into cells, e.g.
PT for treating infection or in gene therapy
XX
PS Example 2; Fig 1; 43pp; French.
XX
CC The invention relates to methods for selecting and identifying a
CC cellular receptor for a virus, by immobilising, on a support, a
CC monoclonal antibody targeted to a viral surface protein that determines
CC attachment of the virus to the receptor. The immobilised antibody is
CC incubated with a hexapeptide expression library and peptides bound to the
CC immobilised antibody are eluted by competitive binding with recombinant
CC fragments of the viral surface protein. In a reverse method, the viral
CC surface protein is immobilised and incubated with the peptide library. In
CC this case, the bound peptides are eluted by competitive binding with the
CC monoclonal antibody. The hexapeptides AAW75278-W75292 represent
CC peptides isolated by biopanning the library with the immobilised
CC monoclonal antibody (MAb) 7A2.7 and eluted using an adenovirus type 5
CC fibre head protein (AAW75267).
CC The methods are used to identify peptides from MHC Class I and/or
CC type III fibronectin proteins that allow or facilitate attachment by
CC adenovirus (Ad) to host cells and/or entry into the cells, and to
CC identify ligands that modulate Ad infection mediated by these peptides,
CC e.g. to treat or prevent Ad infections or to facilitate infection by Ad
CC gene therapy vectors used to treat genetic diseases, acquired immune
CC deficiency syndrome or cancer.

SQ Sequence 6 AA;

Query Match 53.1%; Score 17; DB 19; Length 6;
 Best Local Similarity 75.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0;

OY 1 FTLT 4
 : 111
 DB 3 yltt 6

RESULT 15

AA07656
 ID AAR07656 standard; protein: 7 AA.

XX AAR07656;

XX 04-JAN-1991 (first entry)

XX Ribonuclease reductase inhibitor #9.

XX mammalian ribonuclease reductase inhibitors; tumour treatment;
 KW psoriasis treatment;

XX Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1..1
 FT /label-N-acetyl-L-phenylalanine
 XX

PN EP383190-A.

PD 22-AUG-1990.

XX 08-FEB-1990; 90EP-0102453.

XX 17-FEB-1989; 89CA-2591372.

XX (BIOM-1) BIO-MEGA INC.

XX Guindon Y, Lavellee P, Rakhit S, Cosentino GP;

DR WPI: 1990-255483/34.

XX Oligopeptide deriva. inhibitors of mammalian ribonucleotide
 PT reductase - for control of neoplasms, tumours, psoriasis, etc..

XX Claim 10; Page 15; 16pp; English.

XX The peptide is
 CC synthesised by standard solid-phase methods. It is a selective
 CC inhibitor of mammalian ribonucleotide reductase and can be used to
 CC control abnormal cell proliferation. The inhibitor has relatively low
 CC toxicity and can be applied topically or systemically
 CC See also AAR06487-R06488, AAR07650-R07655 and AAR07657.

XX Sequence 7 AA;

Query Match 53.1%; Score 17; DB 11; Length 7;

Best Local Similarity 60.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTLT 5
 : 111
 DB 1 ftdv 5

Search completed: July 15, 2002, 13:25:24
 Job time: 1455 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:18:00 ; Search time 40.89 Seconds
(without alignments)
4.181 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTFTISS 7

Scoring table: BLOSUM62

Searched: Gap0 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 39160

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56.2	7	4	US-09-187-859-637	Sequence 637, App
2	53.1	5	3	US-08-757-177-16	Sequence 16, Appl
3	53.1	7	2	US-08-739-401A-6	Sequence 6, Appl
4	50.0	4	4	US-08-860-904-9	Sequence 9, Appl
5	50.0	6	1	US-08-252-995D-7	Sequence 7, Appl
6	50.0	6	2	US-08-478-386A-46	Sequence 46, Appl
7	50.0	6	2	US-08-292-597-46	Sequence 46, Appl
8	50.0	6	2	US-08-388-653-46	Sequence 46, Appl
9	50.0	6	2	US-08-473-985-46	Sequence 46, Appl
10	50.0	6	2	US-08-834-108-7	Sequence 7, Appl
11	50.0	6	2	US-08-483-898-46	Sequence 46, Appl
12	50.0	6	3	US-09-087-716-46	Sequence 46, Appl
13	50.0	6	3	US-09-157-753-46	Sequence 46, Appl
14	50.0	6	3	US-09-157-230-46	Sequence 46, Appl
15	50.0	6	3	US-09-087-811-46	Sequence 46, Appl
16	50.0	6	3	US-09-156-855-46	Sequence 46, Appl
17	50.0	6	4	US-09-158-010-46	Sequence 46, Appl
18	50.0	6	4	US-09-087-647-46	Sequence 46, Appl
19	50.0	6	4	US-09-302-629-46	Sequence 46, Appl
20	50.0	7	1	US-08-136-743B-55	Sequence 55, Appl
21	50.0	7	2	US-08-177-109A-7	Sequence 7, Appl
22	50.0	7	2	US-08-687-706-7	Sequence 28, Appl
23	50.0	7	3	US-09-040-216-28	Sequence 63, Appl
24	46.9	5	1	US-08-136-743B-63	Sequence 22, Appl
25	46.9	5	2	US-08-667-001-22	Sequence 55, Appl
26	46.9	5	3	US-09-040-216-59	Sequence 17, Appl
27	46.9	5	4	US-08-591-632-17	

28	15	46.9	5	4	US-08-591-632-23	Sequence 23, Appl
29	15	46.9	5	4	US-08-591-632-26	Sequence 26, Appl
30	15	46.9	6	1	US-08-136-743B-62	Sequence 62, Appl
31	15	46.9	6	1	US-08-297-731-4	Sequence 4, Appl
32	15	46.9	6	1	US-08-290-448A-41	Sequence 41, Appl
33	15	46.9	6	1	US-08-290-448A-41	Sequence 41, Appl
34	15	46.9	6	1	US-08-175-069A-41	Sequence 41, Appl
35	15	46.9	6	2	US-08-622-720A-8	Sequence 8, Appl
36	15	46.9	6	3	US-09-040-216-54	Sequence 54, Appl
37	15	46.9	6	3	US-08-297-395-3	Sequence 3, Appl
38	15	46.9	6	3	US-09-146-675-6	Sequence 6, Appl
39	15	46.9	6	4	US-09-623-618B-6	Sequence 4, Appl
40	15	46.9	6	4	US-08-461-939B-41	Sequence 41, Appl
41	15	46.9	6	4	US-08-464-000-41	Sequence 41, Appl
42	15	46.9	6	4	US-09-187-859-650	Sequence 650, Appl
43	15	46.9	6	5	PCT-US95-10783-4	Sequence 4, Appl
44	15	46.9	7	1	US-08-136-743B-6	Sequence 6, Appl
45	15	46.9	7	1	US-08-136-743B-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-187-859-637
Sequence 637, Application US/09187859A
Patent No. 6358920
GENERAL INFORMATION:
APPLICANT: Blaschuk, Orest W.
APPLICANT: Gour, Barbara J.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
FILE REFERENCE: 100086, 407C1
CURRENT APPLICATION NUMBER: US/09/187, 859A
CURRENT FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 4052
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 637
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Representative linear modulating agent based on
OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
OTHER INFORMATION: recognition sequence
US-09-187-859-637

Query Match 56.2%; Score 18; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FTFTISS 7
Db 1 FTIDSS 7

RESULT 2
US-08-757-177-16
Sequence 16, Application US/08757177
Patent No. 6071718
GENERAL INFORMATION:
APPLICANT: MUKERJI, PRADIP
APPLICANT: HARDS, ROBERT G.
APPLICANT: THURMOND, JENNIFER M.
APPLICANT: LEONARD, AMANDA EUN-YEONG
TITLE OF INVENTION: METHODS OF PRODUCING A RECOMBINANT PROTEIN
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Abbott Laboratories
STREET: One Abbott Park Road
CITY: Abbott Park
STATE: Illinois

```
COUNTRY: USA
ZIP: 60064
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,177
FILING DATE: 27-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6004.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-935-1729
TELEFAX: 848-938-2623
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-757-177-16

Query Match
Best Local Similarity 53.1%; Score 17; DB 3; Length 5;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TISS 7
1111
DB 2 TISS 5

RESULT 3
US-08-739-401A-6
Sequence 6, Application US/08739401A
Patent No. 5837461
GENERAL INFORMATION:
APPLICANT: Neltz, Maureen E.
APPLICANT: Neltz, John F.
TITLE OF INVENTION: DETECTION OF CONE-PHOTORECEPTOR-BASED
TITLE OF INVENTION: VISION DISORDERS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,401A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Jean C.
REGISTRATION NUMBER: 35,433
REFERENCE/DOCKET NUMBER: 65005.91151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5709
TELEFAX: (414) 271-3552
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
```

```
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-739-401A-6

Query Match
Best Local Similarity 53.1%; Score 17; DB 2; Length 7;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PTLT 4
1111
DB 2 PTLT 5

RESULT 4
US-08-860-904-9
Sequence 9, Application US/08860904
Patent No. 6294654
GENERAL INFORMATION:
APPLICANT: Sandile, Inger
APPLICANT: Bogen, Bjarte
APPLICANT: Fossum, Sigbjorn
TITLE OF INVENTION: A Modified Immunoglobulin Molecule
TITLE OF INVENTION: Incorporating an Antigen in a No. 6294654-CDR
FILE REFERENCE: 9914-1
CURRENT APPLICATION NUMBER: US/08/860,904
CURRENT FILING DATE: 1997-09-29
EARLIER APPLICATION NUMBER: PCT/GB96/00116
EARLIER FILING DATE: 1996-01-19
EARLIER APPLICATION NUMBER: GB 9501079.9
EARLIER FILING DATE: 1995-01-19
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 9
LENGTH: 4
TYPE: PRT
ORGANISM: Mus sp.
US-08-860-904-9

Query Match
Best Local Similarity 50.0%; Score 16; DB 4; Length 4;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TISS 7
1111
DB 1 TVSS 4

RESULT 5
US-08-252-995D-7
Sequence 7, Application US/08252995D
Patent No. 5650501
GENERAL INFORMATION:
APPLICANT: Dennis, James W
APPLICANT: Heffernan, Mike
APPLICANT: Fode, Carol
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BEREKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/252,995D
FILING DATE: 02-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurdgyk, Linda M
REGISTRATION NUMBER: 34,972
REFERENCE/DOCKET NUMBER: 3153-96
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-252-995D-7

Query Match 50.0%; Score 16; DB 1; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 7
11:1
Db 2 LTLSN 6

RESULT 6
US-08-478-386A-46
Sequence 46, Application US/08478386A
Patent No. 5830462
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
TITLE OF INVENTION: GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,386A
FILING DATE: 07/JUN/1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-114A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-386A-46

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 6
11:1
Db 3 LTVS 6

RESULT 7
US-08-292-597-46
Sequence 46, Application US/08292597
Patent No. 5834266
GENERAL INFORMATION:
APPLICANT: Gerald R. Crabtree
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: Regulated Apoptosis
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,597
FILING DATE: 18/AUG/1994
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-108A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-292-597-46

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 6
11:1
Db 3 LTVS 6

RESULT 8
US-08-388-653-46
Sequence 46, Application US/08388653

Patent No. 5869337
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
TITLE OF INVENTION: GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,653
FILING DATE: 14-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,386
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-114A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ. ID NO.: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-388-653-46

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 6
DB 3 LTVS 6

RESULT 9
US-08-473-985-46
Sequence 46, Application US/08473985
Patent No. 5871753
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
APPLICANT: Ho, Steffan
TITLE OF INVENTION: Regulated Transcription of Targeted Genes and
TITLE OF INVENTION: Other Biological Events
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California

COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,985
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/179,748
FILING DATE: 07-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-SU 9863
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO.: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-473-985-46

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 6
DB 3 LTVS 6

RESULT 10
US-08-834-108-7
Sequence 7, Application US/08834108
Patent No. 5976893
GENERAL INFORMATION:
APPLICANT: Dennis, James M.
APPLICANT: Helferman, Mike
APPLICANT: Fode, Carol
TITLE OF INVENTION: NOVEL SERINE/THREONINE KINASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,108
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Kurodyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ. ID NO.: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mus musculus
US-08-834-108-7

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTIS 7
11:1
DB 2 LTLS 6

RESULT 11
US-08-483-898-46
Sequence 46, Application US/08483898
Patent No. 5994313
GENERAL INFORMATION:
APPLICANT: Gerald R. Crabtree
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: Regulated Apoptosis
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,898
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/292,597
FILING DATE: 18-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fligg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-108A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-898-46

Query Match 50.0%; Score 16; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 3 LTIS 6
11:1

DB 3 LTVS 6

RESULT 12
US-09-087-716-46
Sequence 46, Application US/09087716
Patent No. 6011018
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/087,716
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/388,653
FILING DATE: 02/14/1995
ATTORNEY/AGENT INFORMATION:
NAME: Fligg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-114A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-087-716-46

Query Match 50.0%; Score 16; DB 3; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 LTIS 6
11:1
DB 3 LTVS 6

RESULT 13
US-09-157-753-46
Sequence 46, Application US/09157753
Patent No. 6043082
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ARIAD Pharmaceuticals, Inc.
;; STREET: 26 Landsdowne Street
;; CITY: Cambridge
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02139
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC/DOS/MS/DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/157,753
;;
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/388,653
;; FILING DATE: 14-FEB-1995
;; APPLICATION NUMBER: US 08/478,386
;; FILING DATE: 07-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fligg, E. Anthony
;; REGISTRATION NUMBER: 27,195
;; REFERENCE/DOCKET NUMBER: 2054-114A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 783-6040
;; TELEFAX: (202) 783-6031
;; INFORMATION FOR SEQ ID NO: 46:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-157-753-46

Query Match 50.0%; Score 16; DB 3; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

OY 3 LTIS 6
||:|
DB 3 LTVS 6

RESULT 14
US-09-157-230-46
; Sequence 46, Application US/09157230
; Patent No. 6046047
; GENERAL INFORMATION:
; APPLICANT: Crabtree, Gerald R.
; APPLICANT: Schreiber, Stuart L.
; APPLICANT: Spencer, David M.
; APPLICANT: Wandless, Thomas J.
; APPLICANT: Belshaw, Peter
; TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
; TITLE OF INVENTION: GENES AND OTHER BIOLOGICAL EVENTS
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARIAD Pharmaceuticals, Inc.
; STREET: 26 Landsdowne Street
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC/DOS/MS/DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,230

;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/478,386
;; FILING DATE: 07/JUN/1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fligg, E. Anthony
;; REGISTRATION NUMBER: 27,195
;; REFERENCE/DOCKET NUMBER: 2054-114A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 783-6040
;; TELEFAX: (202) 783-6031
;; INFORMATION FOR SEQ ID NO: 46:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-09-157-230-46

Query Match 50.0%; Score 16; DB 3; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

OY 3 LTIS 6
||:|
DB 3 LTVS 6

RESULT 15
US-09-087-811-46
; Sequence 46, Application US/09087811
; Patent No. 6054436
; GENERAL INFORMATION:
; APPLICANT: Gerald R. Crabtree
; APPLICANT: Schreiber, Stuart L.
; APPLICANT: Spencer, David M.
; APPLICANT: Wandless, Thomas J.
; APPLICANT: Belshaw, Peter
; TITLE OF INVENTION: Regulated Apoptosis
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARIAD Pharmaceuticals, Inc.
; STREET: 26 Landsdowne Street
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC/DOS/MS/DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/087,811
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/292,597
; FILING DATE: 18-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fligg, E. Anthony
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 2054-108A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 783-6040
; TELEFAX: (202) 783-6031
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-09-087-811-46

Query Match 50.0%; Score 16; DB 3; length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LTIS 6
||:|
Db 3 LTVS 6

Search completed: July 15, 2002, 13:26:12
Job time: 492 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:27:07 ; Search time 49.13 Seconds

(without alignments)
13.691 Million cell updates/sec

Title: US-09-712-819A-2

Perfect score: 33

Sequence: 1 TRFSKL 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_71:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	36.4	6	2	A61411
2	12	36.4	6	2	I59142
3	12	36.4	7	2	A30812
4	10	30.3	4	2	I40505
5	10	30.3	5	2	PT0295
6	10	30.3	6	2	A61068
7	10	30.3	6	2	I49421
8	10	30.3	6	2	PD0028
9	10	30.3	7	2	JN0859
10	10	30.3	7	2	A25269
11	10	30.3	7	2	PX0008
12	10	30.3	7	2	PD0029
13	10	30.3	7	4	I55382
14	9	27.3	4	2	T46627
15	9	27.3	4	2	J01273
16	9	27.3	4	2	E44823
17	9	27.3	4	2	PT0534
18	9	27.3	5	2	C41225
19	9	27.3	5	2	I40702
20	9	27.3	5	2	S11127
21	9	27.3	5	2	S62883
22	9	27.3	6	2	A60986
23	9	27.3	6	2	A60494
24	9	27.3	6	2	I37263
25	9	27.3	6	2	I46126
26	9	27.3	6	2	B33932
27	9	27.3	6	2	I49424
28	9	27.3	7	2	B39127
29	9	27.3	7	2	I40504

30	9	27.3	7	2	S68004	hucollin, 75K chain
31	8	24.2	4	2	PT0645	T-cell receptor be
32	8	24.2	5	2	A44692	fulcin - giant Af
33	8	24.2	5	2	PT0278	Ig heavy chain CRD
34	8	24.2	5	2	PT0624	T-cell receptor be
35	8	24.2	5	2	PT0625	T-cell receptor be
36	8	24.2	5	2	PT0713	T-cell receptor be
37	8	24.2	6	2	I51317	bHh transcription
38	8	24.2	6	2	PT0519	T-cell receptor be
39	8	24.2	6	2	PT0621	T-cell receptor be
40	8	24.2	6	2	PT0730	T-cell receptor be
41	8	24.2	7	2	PH1408	Ig heavy chain V r
42	8	24.2	7	2	S29735	polyphosphate--glu
43	8	24.2	7	2	A34818	vicillin 72K chain
44	8	24.2	7	2	PN0150	omega-glialadin 1'
45	8	24.2	7	2	S33567	tubulin beta-3 cha

ALIGNMENTS

RESULT 1
A61411
ameletin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
C:Accession: A61411
R:Burzyski, S.R.
Anal. Blochem. 70, 359-365, 1976
A>Title: Sequential analysis in subnanomolar amounts of peptides. Determination of th
A:Reference number: A61411; MUID:76182447
A:Accession: A61411
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-6 <BUR>
C:Superfamily: unassigned animal peptides
C:Keywords: pyroglutamic acid
F1/Modified site: Pyroglutamate carboxylic acid (Gln) #status experimental

Query Match 36.4%; Score 12; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 ESK 6
DB 4 YSK 6
RESULT 2
I59142
platelet-derived growth factor B chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I59142
R:Peck, M.; Gazit, A.; Arnstein, P.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 86, 2693-2697, 1989
A>Title: Generation of fibrosarcomas in vivo by a retrovirus that expressed the norma
A:Reference number: I59142; MUID:89202393
A:Accession: I59142
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-6 <RES>
A:Cross-references: GB:M26180; NID:g516624; PIDN:AAA39905.1; PID:g516625

Query Match 36.4%; Score 12; DB 2; Length 6;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 IRSKL 7
DB 2 VFAVL 6

RESULT 3
A:0812
sex pheromone ccf10 - Enterococcus faecalis
C:Species: Enterococcus faecalis
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 18-Jun-1993
C:Accession: A30812
R:Morl, M.; Sakagami, Y.; Ishii, Y.; Isogai, A.; Kikada, C.; Fujino, M.; Adachi, J.C.; Du
J. Biol. Chem. 263, 14574-14578, 1988
A:Title: Structure of ccf10, a peptide sex pheromone which induces conjugative transfer
A:Reference number: A30812; MUID:99008313
A:Accession: A30812
A:Molecule type: protein
A:Residues: 1-7 <MOR>

Query Match 36.4%; Score 12; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIP 4
1 : 1
DB 3 TLVF 6

RESULT 4
140505
hypothetical protein 3 (4 aa) - Bacillus stearothermophilus
C:Species: Bacillus stearothermophilus
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
C:Accession: 140505
R:Waye, M.M.; Winter, G.
Eur. J. Biochem. 158, 505-510, 1986
A:Title: A transcription terminator in the 5' non-coding region of the tyrosyl tRNA synth
A:Reference number: 140503; MUID:86274732
A:Accession: 140505
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4 <RES>
A:Cross-references: EMBL:X04193; NID:g40233; PIDN:CAA27783.1; PID:g580944

Query Match 30.3%; Score 10; DB 2; Length 4;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 IFSK 6
1 : 1
DB 1 MLSK 4

RESULT 5
PT0295
Ig heavy chain CRD3 region (clone 5-91) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C:Accession: PT0295
R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A:Reference number: PT0222; MUID:91108337
A:Accession: PT0295
A:Molecule type: DNA
A:Residues: 1-5 <YAM>
A:Experimental source: B lymphocyte
C:Keywords: heterotetramer; immunoglobulin

Query Match 30.3%; Score 10; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TR 2
1 : 1
DB 3 TR 4

RESULT 6
A61068
locustakinin - migratory locust
C:Species: Locusta migratoria (migratory locust)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
C:Accession: A61068
R:Schodde, D.; Holman, G.M.; Proost, P.; Van Damme, J.; Hayes, T.K.; De Loof, A.
Regul. Pept. 37, 49-57, 1992
A:Title: Locustakinin, a novel myotropic peptide from Locusta migratoria, isolation,
A:Reference number: A61068; MUID:92262851
A:Accession: A61068
A:Molecule type: protein
A:Residues: 1-6 <SCH>
C:Keywords: amidated carboxyl end; cephalomyotropic peptide; neuropeptide
F:6/modified site: amidated carboxyl end (gly) #status experimental

Query Match 30.3%; Score 10; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
1 : 1
DB 2 FS 3

RESULT 7
149421
laminin B1 - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: 149421
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maeraki, Y.; Nadeau,
Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: 148934; MUID:94319082
A:Accession: 149421
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6 <RES>
A:Cross-references: EMBL:U05736; NID:g497073; PIDN:AAB60477.1; PID:g642829

Query Match 30.3%; Score 10; DB 2; Length 6;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IFS 5
1 : 1
DB 1 VYS 3

RESULT 8
PD0028
pev-kinin 2 - penaeid shrimp (Penaeus vannamei) (fragment)
C:Species: Penaeus vannamei
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000
C:Accession: PD0028
R:Nieto, J.; Vealeert, D.; Derua, R.; Waelkens, E.; Cersiaens, A.; Coast, G.; Devree
Biochem. Biophys. Res. Commun. 248, 406-411, 1998
A:Title: Identification of one tachykinin- and two kinin-related peptides in the brain
A:Reference number: PD0027; MUID:98342103
A:Accession: PD0028
A:Molecule type: protein
A:Residues: 1-6 <NIE>
C:Comment: This peptide belongs to myotropic neuropeptides.

Query Match 30.3%; Score 10; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
 11
 DB 2 FS 3

RESULT 9

JN0859
 peptidyl-dipeptidase A inhibitory peptide C105 - striped bonito

C:Species: Sarda orientalis (striped bonito)
 C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999

C:Accession: JN0859

R:Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.

Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993

A:Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory pe

A:Reference number: JN0859; MUID:94080036

A:Accession: JN0859

A:Molecule type: protein

A:Residues: 1-7 <MAT>

A:Experimental source: intestine

C:Comment: The carboxyl-terminus is essential for the protein's expression of angiotensi

C:Superfamily: bradykinin-potentiating peptide

C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 30.3%; Score 10; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 SKL 7
 11
 DB 3 AKL 5

RESULT 10

A25269
 sex pheromone CAM373 - Enterococcus faecalis

N:Alternate names: clumping-inducing agent (CIA)

C:Species: Enterococcus faecalis

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 18-Jun-1993

C:Accession: A25269

R:Morit, M.; Tanaka, H.; Sakagami, Y.; Isogai, A.; Fujino, M.; Kitada, C.; White, B.A.; A

FBRS Lett. 206, 68-72, 1986

A:Title: Isolation and structure of the Streptococcus faecalis sex pheromone, CAM373.

A:Reference number: A25269; MUID:87005252

A:Accession: A25269

A:Molecule type: protein

A:Residues: 1-7 <MOR>

Query Match 30.3%; Score 10; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 IF 4
 11
 DB 2 IF 3

RESULT 11

PX0008
 glucuronosyltransferase (EC 2.4.1.17), hepatic - rat (fragment)

N:Alternate names: UDP-glucuronosyltransferase

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 07-Feb-1997

C:Accession: PX0008

R:Yokota, H.; Yuasa, A.; Sato, R.

J. Biochem. 104, 531-536, 1988

A:Title: Purification and properties of a form of UDP-glucuronosyltransferase from liver m

A:Accession: PX0008
 A:Molecule type: protein
 A:Residues: 1-7 <YOK>
 C:Keywords: glucosyltransferase; hexosyltransferase; liver

Query Match 30.3%; Score 10; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 SKL 7
 11
 DB 1 TKL 3

RESULT 12

PD0029
 pex-kinin 1 - penaeid shrimp (Penaeus vannamei) (fragment)

C:Species: Penaeus vannamei

C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000

C:Accession: PD0029

R:Mieto, J.; Veelaert, D.; Deruy, R.; Maelkens, E.; Cerstiaens, A.; Coast, G.; Devree

Biochem. Biophys. Res. Commun. 248, 406-411, 1998

A:Title: Identification of one tachykinin- and two kinin-related peptides in the brat

A:Reference number: PD0027; MUID:98342103

A:Accession: PD0029

A:Molecule type: protein

A:Residues: 1-7 <NIE>

C:Comment: This peptide belongs to myotropic neuropeptides.

Query Match 30.3%; Score 10; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
 11
 DB 3 FS 4

RESULT 13

I55382
 hypothetical peptide PAII promoter region - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 16-Apr-1999 #sequence_revision 16-Apr-1999 #text_change 20-Apr-2000

C:Accession: I55382

R:Lawson, S.J.; Wiman, B.; Hamsten, A.; Green, F.; Humphries, S.; Henney, A.M.

J. Biol. Chem. 268, 10739-10745, 1993

A:Title: The two allele sequences of a common polymorphism in the promoter of the pla

A:Reference number: I55382; MUID:93266509

A:Accession: I55382

A>Status: translation not shown; translated from GB/EWML/DBRJ

A:Molecule type: DNA

A:Residues: 1-7 <DNA>

A:Cross-references: GB:M1557; NID:q190020; PIDN:AA60110.1; PID:q190021

C:Comment: This is the hypothetical translation of a sequence from the PAII gene prom

C:Genetics:

A:Gene: GDB:PAII

A:Cross-references: GDB:120297; OMIM:173360

A:Map position: 7q21.3-7q22

Query Match 30.3%; Score 10; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TR 2
 11
 DB 2 TR 3

RESULT 14

T46627

hypothetical protein c4 - loblolly pine
 C:Species: Pinus taeda (loblolly pine)
 C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 C:Accession: T46627
 R:Chang, S.; Purysa, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.
 submitted to the EMBL Data Library, July 1995
 A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is dc
 A:Reference number: 223105
 A:Accession: T46627
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-4 <CHA>
 A:Cross-references: EMBL:U01309; NID:g974285; PID:g974292
 A:Experimental source: strain s6PT2xs6PT3; 8 month seedlings

Query Match 27.3%; Score 9; DB 2; Length 4;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KL 7
 11
 DB 2 KL 3

RESULT 15

J01273
 neuropeptide Antho-Kamide - sea anemone (Anthopleura elegantissima)
 C:Species: Anthopleura elegantissima
 C>Date: 31-Mar-1992 #sequence_revision 04-Dec-1992 #text_change 08-Dec-1995
 C:Accession: J01273
 R:Notackker, H.P.; Rinehart, K.L.; Grimmelikhuijzen, C.J.P.
 Biochem. Biophys. Res. Commun. 179, 1205-1211, 1991
 A:Title: Isolation of L-3-phenylacetyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a novel neuropep
 A:Reference number: J01273; MUID:92028852
 A:Accession: J01273
 A:Molecule type: protein
 A:Residues: 1-4 <NOT>
 C:Comment: The carboxyl-terminal amide probably arises from cleavage of a following glyc
 C:Keywords: amidated carboxyl end; neuropeptide; phenylacetylation
 F:1/Modified site: L-3-phenylacetic acid (Phe) #status experimental
 F:4/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 27.3%; Score 9; DB 2; Length 4;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 FSK 6
 11
 DB 1 FFK 3

Search completed: July 15, 2002, 13:27:08
 Job time: 438 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:38:10 ; Search time 24.49 Seconds

(without alignments)
11.067 Million cell updates/sec

Title: US-09-712-819A-2

Perfect score: 33

Sequence: 1 TRIFSKL 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB ID	Description
1	12	36.4	7	1	CCP1_ENTFA
2	10	30.3	5	1	UC22_MAIZE
3	10	30.3	6	1	LOK1_LOCOMI
4	10	30.3	7	1	CIA_ENTFA
5	9	27.3	4	1	FFKA_ANTEL
6	9	27.3	4	1	FTRI_ANTEL
7	9	27.3	5	1	PSK_DAUCA
8	9	27.3	7	1	UN06_PINPS
9	8	24.2	7	1	FAR1_PROCL
10	8	24.2	7	1	FAR2_PROCL
11	7	21.2	4	1	ACH1_ACHFU
12	7	21.2	4	1	FMRF_MANCN
13	7	21.2	4	1	TUFT_HUMAN
14	7	21.2	5	1	AL14_CARMA
15	7	21.2	5	1	RE11_LITRU
16	7	21.2	5	1	RE21_LITRU
17	7	21.2	5	1	RE31_LITRU
18	7	21.2	5	1	RE32_LITRU
19	7	21.2	6	1	CIF2_MYTED
20	7	21.2	6	1	CIF2_MYTED
21	7	21.2	6	1	TM06_SARPO
22	7	21.2	7	1	AL17_CYPDO
23	7	21.2	7	1	CARP_MYTED
24	7	21.2	7	1	E105_LITRU
25	7	21.2	7	1	FARB_CAIYO
26	7	21.2	7	1	HY7_PIG
27	7	21.2	7	1	LANC_CARUT
28	7	21.2	7	1	TY51_LITRU
29	6	18.2	4	1	DCMS_PSECH
30	6	18.2	4	1	FAR3_HIRME
31	6	18.2	4	1	FAR4_HIRME
32	6	18.2	4	1	FLRF_HIRME
33	6	18.2	4	1	FLRN_ANTEL

34	6	18.2	4	1	CCP1_OCTMT	P58648 octopus min
35	6	18.2	5	1	E103_LITRU	P82099 litoria rub
36	6	18.2	5	1	E104_LITRU	P83100 litoria rub
37	6	18.2	5	1	FARP_ARTRR	P41853 artiposthi
38	6	18.2	5	1	PAP2_PARMA	P81864 pardachilus
39	6	18.2	5	1	SUGA_ACHDO	P19991 acheta dome
40	6	18.2	5	1	TPIS_CANFA	P54714 canis famli
41	6	18.2	6	1	E101_LITRU	P82096 litoria rub
42	6	18.2	6	1	FARP_MONEX	P81806 monleza ex
43	6	18.2	7	1	ALL2_CARMA	P81805 carclnus ma
44	6	18.2	7	1	ALL3_CARMA	P81806 carclnus ma
45	6	18.2	7	1	ALL4_CARMA	P81807 carclnus ma

ALIGNMENTS

```

RESULT 1
ID CCP1_ENTFA STANDARD: PRT: 7 AA.
AC P20104:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CCP10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OX Enterococcus
NCBI_TaxID=1351;
RN [1]
RP MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adachi J.C., Dunnay G.M., Suzuki A.;
RT Structure of CCP10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, pCCF10.
RL J. Biol. Chem. 263:14574-14578(1988).
CC -1- FUNCTION: CCP10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC HEMOLYSIN PLASMID PCF10.
DR PIR: A30812; A30812.
SQ
SEQUENCE 7 AA; 790 MW; 72C9D2C731B2C740 CRC64;

```

Query Match 36.4%; Score 12; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRIF 4
DB 3 TLVF 6

```

RESULT 2
ID UC22_MAIZE STANDARD: PRT: 5 AA.
AC P80628:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1996 (Rel. 38, Last annotation update)
DE Unknown protein from Zm-page of etiolated coleoptile (Spot 474)
DE (Fragment).
OS Zea mays (maize).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Perollet J.-C., Zlivy M., de Vienne D.;

```

RT "The maize two dimensional gel protein database: towards an integrated
 genome analysis program."
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PT OF THIS UNKNOWN
 DR PROTEIN IS: 6.1, ITS MW IS: 30.4 kDa.
 DR MaizeDB: P80628; COLEOPTILE.
 DR MaizeDB: 123954; "-."
 RT NON_TER 1 1
 FT 5 5
 SQ SEQUENCE 5 AA: 654 MW: 72CB19C9C0300000 CRC64;

Query Match 30.3%; Score 10; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 IF 4
 DB 1 IF 2

RESULT 3
 LOK1_LOCM1 STANDARD; PRT; 6 AA.
 AC P41491;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Locustakinin I.
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caellifera;
 OX Acridomorpha; Acridoidea; Acrididae; Locusta.
 RN NCB1_Taxid=7004;
 RP SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RX MEDLINE=92262851; PubMed=1585017;
 RA Schoofs L., Holman G.M., Proost P., van Damme J., Hayes T.K.,
 RA de Loof A.;
 RT "Locustakinin, a novel myotropic peptide from Locusta migratoria,
 RT isolation, primary structure and synthesis.";
 RL Regul. Pept. 37:45-57(1992).
 CC -1- FUNCTION: MYOTROPIC PEPTIDE. MAY BE IMPORTANT IN THE STIMULATION
 CC OF ION TRANSPORT AND INHIBITION OF DIURETIC ACTIVITY IN MALPIGHIAN
 CC TUBULES.
 DR PIR: A61068; A61068.
 KW Neuropeptide; Amidation.
 FT MOD_RES 6 6
 SQ SEQUENCE 6 AA: 654 MW: 686365A5B9CDB000 CRC64;

Query Match 30.3%; Score 10; DB 1; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
 DB 2 FS 3

RESULT 4
 C1A_ENTFA STANDARD; PRT; 7 AA.
 AC P11932;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone CAM373 (clumping-inducing agent) (C1A).
 OS Enterococcus faecalis (Streptococcus faecalis).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 OX NCB1_Taxid=1351;

RN [1]
 RP SEQUENCE.
 RX MEDLINE=87005252; PubMed=3093276;
 RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
 RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
 RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
 RT CAM373.";
 RL FEBS Lett. 206:69-72(1986).
 CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
 CC HARBORING PAM373.
 CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
 CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
 CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
 DR PIR: A25269; A25269.
 KW Pheromone.
 SQ SEQUENCE 7 AA: 734 MW: 75BDD72059C05D80 CRC64;

Query Match 30.3%; Score 10; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 IF 4
 DB 2 IF 3

RESULT 5
 FFKA_ANTEL STANDARD; PRT; 4 AA.
 AC P58705;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Antho-Kamide.
 OS Anthopleura elegantissima (Sea anemone).
 OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
 CC Nyanthaeae; Actiniidae; Anthopleura.
 OX NCB1_Taxid=6110;
 RN [1]
 RP SEQUENCE.
 RX PubMed=1681803;
 RA Nohacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
 RT "Isolation of L-3-phenylalanyl-phe-Lys-Ala-NH2 (Antho-Kamide), a
 RT novel neuropeptide from sea anemones.";
 RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
 RN [2]
 RP FUNCTION.
 RX PubMed=8397415;
 RA McFarlane I.D., Hudman D., Nohacker H.-P., Grimmelikhuijzen C.J.P.;
 RT "The expansion behaviour of sea anemones may be coordinated by two
 RT inhibitory neuropeptides, Antho-Kamide and Antho-Ramide.";
 RL Proc. R. Soc. Lond. B, Biol. Sci. 253:183-188(1993).
 CC -1- FUNCTION: Inhibits spontaneous contractions in several muscle
 CC groups. May be involved in the expansion phase of feeding
 CC behaviour in sea anemones.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: Neuron-specific.
 KW Neuropeptide; Amidation.
 FT MOD_RES 1 1
 FT MOD_RES 4 4
 SQ SEQUENCE 4 AA: 512 MW: 6DD339C9A0000000 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 4;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 FSK 6
 DB 1 FFK 3

```

RESULT 6
FYRI ANTEL STANDARD; PRT; 4 AA.
ID P58706;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Antho-Riamide I (Contains: Antho-Riamide II).
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nyantheae; Actiniidae; Anthopleura.
NCBI_TaxID=6110;
RX PUBMED=1821096;
RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA Gimmelikhuijzen C.J.P.;
RT *Isolation of two novel neuropeptides from sea anemones: the unusual,
RT biologically active L-3-phenylalanyl-Tyr-Arg-Ile-NH2 and its
RT des-phenylalanyl fragment Tyr-Arg-Ile-NH2.
RL Peptides 12:1165-1173(1991).
RN [2]
RP FUNCTION.
RX PUBMED=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Gimmelikhuijzen C.J.P.;
RT *The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Riamide and Antho-Riamide.
Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -I- FUNCTION: inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC Neuropeptide; Amidation.
KM CHAIN 1 4 ANTHO-RIAMIDE I.
FT MOD_RES 2 4 ANTHO-RIAMIDE II.
FT MOD_RES 1 1 L-3-PHENYLALANYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 598 MW; 60441B39A0000000 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RI 3
   1 1
   2 2
Db 3 RI 4

RESULT 7
PSK DAUCA STANDARD; PRT; 5 AA.
ID P58261;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-beta)].
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RX STRAIN=cv. US-Harumakigosun;
RX MEDLINE=20212743; PubMed=10750705;
RA Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
RA Kanada H., Sakagami Y.;
RT "A secreted peptide growth factor, phytosulfokine, acting as a
RT stimulatory factor of carrot somatic embryo formation."
RT Plant Cell Physiol. 41:27-32(2000).

```

```

CC -I- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
CC CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
CC EMBRYOS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
KM Growth factor; Sulfation.
FT PEPTIDE 1 4 PHYTOSULFOKINE-BETA.
FT MOD_RES 1 1 SULFATION.
FT MOD_RES 3 3 SULFATION.
SQ SEQUENCE 5 AA; 687 MW; 76C1B504B3000000 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 5;
Best Local Similarity 25.0%; Pred. No. 1e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFSK 6
   1 1
   2 2
   3 3
   4 4
   5 5
   6 6
   7 7
   8 8
   9 9
   10 10
   11 11
   12 12
   13 13
   14 14
   15 15
   16 16
   17 17
   18 18
   19 19
   20 20
   21 21
   22 22
   23 23
   24 24
   25 25
   26 26
   27 27
   28 28
   29 29
   30 30
   31 31
   32 32
   33 33
   34 34
   35 35
   36 36
   37 37
   38 38
   39 39
   40 40
   41 41
   42 42
   43 43
   44 44
   45 45
   46 46
   47 47
   48 48
   49 49
   50 50
   51 51
   52 52
   53 53
   54 54
   55 55
   56 56
   57 57
   58 58
   59 59
   60 60
   61 61
   62 62
   63 63
   64 64
   65 65
   66 66
   67 67
   68 68
   69 69
   70 70
   71 71
   72 72
   73 73
   74 74
   75 75
   76 76
   77 77
   78 78
   79 79
   80 80
   81 81
   82 82
   83 83
   84 84
   85 85
   86 86
   87 87
   88 88
   89 89
   90 90
   91 91
   92 92
   93 93
   94 94
   95 95
   96 96
   97 97
   98 98
   99 99
   100 100
   101 101
   102 102
   103 103
   104 104
   105 105
   106 106
   107 107
   108 108
   109 109
   110 110
   111 111
   112 112
   113 113
   114 114
   115 115
   116 116
   117 117
   118 118
   119 119
   120 120
   121 121
   122 122
   123 123
   124 124
   125 125
   126 126
   127 127
   128 128
   129 129
   130 130
   131 131
   132 132
   133 133
   134 134
   135 135
   136 136
   137 137
   138 138
   139 139
   140 140
   141 141
   142 142
   143 143
   144 144
   145 145
   146 146
   147 147
   148 148
   149 149
   150 150
   151 151
   152 152
   153 153
   154 154
   155 155
   156 156
   157 157
   158 158
   159 159
   160 160
   161 161
   162 162
   163 163
   164 164
   165 165
   166 166
   167 167
   168 168
   169 169
   170 170
   171 171
   172 172
   173 173
   174 174
   175 175
   176 176
   177 177
   178 178
   179 179
   180 180
   181 181
   182 182
   183 183
   184 184
   185 185
   186 186
   187 187
   188 188
   189 189
   190 190
   191 191
   192 192
   193 193
   194 194
   195 195
   196 196
   197 197
   198 198
   199 199
   200 200
   201 201
   202 202
   203 203
   204 204
   205 205
   206 206
   207 207
   208 208
   209 209
   210 210
   211 211
   212 212
   213 213
   214 214
   215 215
   216 216
   217 217
   218 218
   219 219
   220 220
   221 221
   222 222
   223 223
   224 224
   225 225
   226 226
   227 227
   228 228
   229 229
   230 230
   231 231
   232 232
   233 233
   234 234
   235 235
   236 236
   237 237
   238 238
   239 239
   240 240
   241 241
   242 242
   243 243
   244 244
   245 245
   246 246
   247 247
   248 248
   249 249
   250 250
   251 251
   252 252
   253 253
   254 254
   255 255
   256 256
   257 257
   258 258
   259 259
   260 260
   261 261
   262 262
   263 263
   264 264
   265 265
   266 266
   267 267
   268 268
   269 269
   270 270
   271 271
   272 272
   273 273
   274 274
   275 275
   276 276
   277 277
   278 278
   279 279
   280 280
   281 281
   282 282
   283 283
   284 284
   285 285
   286 286
   287 287
   288 288
   289 289
   290 290
   291 291
   292 292
   293 293
   294 294
   295 295
   296 296
   297 297
   298 298
   299 299
   300 300
   301 301
   302 302
   303 303
   304 304
   305 305
   306 306
   307 307
   308 308
   309 309
   310 310
   311 311
   312 312
   313 313
   314 314
   315 315
   316 316
   317 317
   318 318
   319 319
   320 320
   321 321
   322 322
   323 323
   324 324
   325 325
   326 326
   327 327
   328 328
   329 329
   330 330
   331 331
   332 332
   333 333
   334 334
   335 335
   336 336
   337 337
   338 338
   339 339
   340 340
   341 341
   342 342
   343 343
   344 344
   345 345
   346 346
   347 347
   348 348
   349 349
   350 350
   351 351
   352 352
   353 353
   354 354
   355 355
   356 356
   357 357
   358 358
   359 359
   360 360
   361 361
   362 362
   363 363
   364 364
   365 365
   366 366
   367 367
   368 368
   369 369
   370 370
   371 371
   372 372
   373 373
   374 374
   375 375
   376 376
   377 377
   378 378
   379 379
   380 380
   381 381
   382 382
   383 383
   384 384
   385 385
   386 386
   387 387
   388 388
   389 389
   390 390
   391 391
   392 392
   393 393
   394 394
   395 395
   396 396
   397 397
   398 398
   399 399
   400 400
   401 401
   402 402
   403 403
   404 404
   405 405
   406 406
   407 407
   408 408
   409 409
   410 410
   411 411
   412 412
   413 413
   414 414
   415 415
   416 416
   417 417
   418 418
   419 419
   420 420
   421 421
   422 422
   423 423
   424 424
   425 425
   426 426
   427 427
   428 428
   429 429
   430 430
   431 431
   432 432
   433 433
   434 434
   435 435
   436 436
   437 437
   438 438
   439 439
   440 440
   441 441
   442 442
   443 443
   444 444
   445 445
   446 446
   447 447
   448 448
   449 449
   450 450
   451 451
   452 452
   453 453
   454 454
   455 455
   456 456
   457 457
   458 458
   459 459
   460 460
   461 461
   462 462
   463 463
   464 464
   465 465
   466 466
   467 467
   468 468
   469 469
   470 470
   471 471
   472 472
   473 473
   474 474
   475 475
   476 476
   477 477
   478 478
   479 479
   480 480
   481 481
   482 482
   483 483
   484 484
   485 485
   486 486
   487 487
   488 488
   489 489
   490 490
   491 491
   492 492
   493 493
   494 494
   495 495
   496 496
   497 497
   498 498
   499 499
   500 500
   501 501
   502 502
   503 503
   504 504
   505 505
   506 506
   507 507
   508 508
   509 509
   510 510
   511 511
   512 512
   513 513
   514 514
   515 515
   516 516
   517 517
   518 518
   519 519
   520 520
   521 521
   522 522
   523 523
   524 524
   525 525
   526 526
   527 527
   528 528
   529 529
   530 530
   531 531
   532 532
   533 533
   534 534
   535 535
   536 536
   537 537
   538 538
   539 539
   540 540
   541 541
   542 542
   543 543
   544 544
   545 545
   546 546
   547 547
   548 548
   549 549
   550 550
   551 551
   552 552
   553 553
   554 554
   555 555
   556 556
   557 557
   558 558
   559 559
   560 560
   561 561
   562 562
   563 563
   564 564
   565 565
   566 566
   567 567
   568 568
   569 569
   570 570
   571 571
   572 572
   573 573
   574 574
   575 575
   576 576
   577 577
   578 578
   579 579
   580 580
   581 581
   582 582
   583 583
   584 584
   585 585
   586 586
   587 587
   588 588
   589 589
   590 590
   591 591
   592 592
   593 593
   594 594
   595 595
   596 596
   597 597
   598 598
   599 599
   600 600
   601 601
   602 602
   603 603
   604 604
   605 605
   606 606
   607 607
   608 608
   609 609
   610 610
   611 611
   612 612
   613 613
   614 614
   615 615
   616 616
   617 617
   618 618
   619 619
   620 620
   621 621
   622 622
   623 623
   624 624
   625 625
   626 626
   627 627
   628 628
   629 629
   630 630
   631 631
   632 632
   633 633
   634 634
   635 635
   636 636
   637 637
   638 638
   639 639
   640 640
   641 641
   642 642
   643 643
   644 644
   645 645
   646 646
   647 647
   648 648
   649 649
   650 650
   651 651
   652 652
   653 653
   654 654
   655 655
   656 656
   657 657
   658 658
   659 659
   660 660
   661 661
   662 662
   663 663
   664 664
   665 665
   666 666
   667 667
   668 668
   669 669
   670 670
   671 671
   672 672
   673 673
   674 674
   675 675
   676 676
   677 677
   678 678
   679 679
   680 680
   681 681
   682 682
   683 683
   684 684
   685 685
   686 686
   687 687
   688 688
   689 689
   690 690
   691 691
   692 692
   693 693
   694 694
   695 695
   696 696
   697 697
   698 698
   699 699
   700 700
   701 701
   702 702
   703 703
   704 704
   705 705
   706 706
   707 707
   708 708
   709 709
   710 710
   711 711
   712 712
   713 713
   714 714
   715 715
   716 716
   717 717
   718 718
   719 719
   720 720
   721 721
   722 722
   723 723
   724 724
   725 725
   726 726
   727 727
   728 728
   729 729
   730 730
   731 731
   732 732
   733 733
   734 734
   735 735
   736 736
   737 737
   738 738
   739 739
   740 740
   741 741
   742 742
   743 743
   744 744
   745 745
   746 746
   747 747
   748 748
   749 749
   750 750
   751 751
   752 752
   753 753
   754 754
   755 755
   756 756
   757 757
   758 758
   759 759
   760 760
   761 761
   762 762
   763 763
   764 764
   765 765
   766 766
   767 767
   768 768
   769 769
   770 770
   771 771
   772 772
   773 773
   774 774
   775 775
   776 776
   777 777
   778 778
   779 779
   780 780
   781 781
   782 782
   783 783
   784 784
   785 785
   786 786
   787 787
   788 788
   789 789
   790 790
   791 791
   792 792
   793 793
   794 794
   795 795
   796 796
   797 797
   798 798
   799 799
   800 800
   801 801
   802 802
   803 803
   804 804
   805 805
   806 806
   807 807
   808 808
   809 809
   810 810
   811 811
   812 812
   813 813
   814 814
   815 815
   816 816
   817 817
   818 818
   819 819
   820 820
   821 821
   822 822
   823 823
   824 824
   825 825
   826 826
   827 827
   828 828
   829 829
   830 830
   831 831
   832 832
   833 833
   834 834
   835 835
   836 836
   837 837
   838 838
   839 839
   840 840
   841 841
   842 842
   843 843
   844 844
   845 845
   846 846
   847 847
   848 848
   849 849
   850 850
   851 851
   852 852
   853 853
   854 854
   855 855
   856 856
   857 857
   858 858
   859 859
   860 860
   861 861
   862 862
   863 863
   864 864
   865 865
   866 866
   867 867
   868 868
   869 869
   870 870
   871 871
   872 872
   873 873
   874 874
   875 875
   876 876
   877 877
   878 878
   879 879
   880 880
   881 881
   882 882
   883 883
   884 884
   885 885
   886 886
   887 887
   888 888
   889 889
   890 890
   891 891
   892 892
   893 893
   894 894
   895 895
   896 896
   897 897
   898 898
   899 899
   900 900
   901 901
   902 902
   903 903
   904 904
   905 905
   906 906
   907 907
   908 908
   909 909
   910 910
   911 911
   912 912
   913 913
   914 914
   915 915
   916 916
   917 917
   918 918
   919 919
   920 920
   921 921
   922 922
   923 923
   924 924
   925 925
   926 926
   927 927
   928 928
   929 929
   930 930
   931 931
   932 932
   933 933
   934 934
   935 935
   936 936
   937 937
   938 938
   939 939
   940 940
   941 941
   942 942
   943 943
   944 944
   945 945
   946 946
   947 947
   948 948
   949 949
   950 950
   951 951
   952 952
   953 953
   954 954
   955 955
   956 956
   957 957
   958 958
   959 959
   960 960
   961 961
   962 962
   963 963
   964 964
   965 965
   966 966
   967 967
   968 968
   969 969
   970 970
   971 971
   972 972
   973 973
   974 974
   975 975
   976 976
   977 977
   978 978
   979 979
   980 980
   981 981
   982 982
   983 983
   984 984
   985 985
   986 986
   987 987
   988 988
   989 989
   990 990
   991 991
   992 992
   993 993
   994 994
   995 995
   996 996
   997 997
   998 998
   999 999
   1000 1000

RESULT 9
FARL PROCL STANDARD; PRT; 7 AA.
ID FARL PROCL
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Cardiac excitatory FMRFamide homolog NFI.
OS Procambrius clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidae;
OC Astacoidea; Cambaridae; Procambrius.
OX NCBI_TaxID=6728;

```

RN [1]
 RP SEQUENCE.
 RC TISSUE-pericardial organs;
 RX MEDLINE-93248032; Pubmed-8387183;
 RA Mercier A.J., Orchard I., Tedrugg V., Skerrett M.;
 RT "Isolation of two FMRFamide-related peptides from crayfish
 pericardial organs.";
 RL Peptides 14:137-143(1993).
 CC -1- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
 CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
 EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide: Amidation.
 FT MOD.RES 7
 SQ SEQUENCE 7 AA; 966 MW; 69D40729C4540420 CRC64;

Query Match 24.2%: Score 8; DB 1; Length 7;
 Best Local Similarity 66.7%: Pred. No. 1e-05; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

Oy 2 RIF 4
 Db 2 RNF 4

RESULT 10
 PAR2_PROCL STANDARD; PRT; 7 AA.
 AC P38498;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Cardioexcitatory FMRFamide homolog DF2.
 OS Procambarus clarkii (Red swamp crayfish).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 CC Astacoidae; Cambaridae; Procambarus.
 OX NCB1_TaxID=6728;
 RN (1)
 RP SEQUENCE.
 RC TISSUE-pericardial organs;
 RX MEDLINE-93248032; Pubmed-8387183;
 RA Mercier A.J., Orchard I., Tedrugg V., Skerrett M.;
 RT "Isolation of two FMRFamide-related peptides from crayfish
 pericardial organs.";
 RL Peptides 14:137-143(1993).
 CC -1- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
 CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
 EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 KW Neuropeptide: Amidation.
 FT MOD.RES 7
 SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540420 CRC64;

Query Match 24.2%: Score 8; DB 1; Length 7;
 Best Local Similarity 66.7%: Pred. No. 1e-05; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

Oy 2 RIF 4
 Db 2 RNF 4

RESULT 11
 ACHI_ACHFU STANDARD; PRT; 4 AA.
 AC P35904;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Achatin-I.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Achatinacea; Achatinidae; Achatina.
 OX NCB1_TaxID=6530;
 RN (1)
 RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.
 RC STRAIN-PERUSSAC; TISSUE-Ganglion;
 RX MEDLINE-89273551; Pubmed-2597281;
 RA Kametani Y., Minakata H., Kenny P.T.M., Iwashita T., Matanabe K.,
 RA Funase K., Sun X.P., Yonagiri A., Kim K.H., Novales-Li P.,
 RA Novales E.T., Kanapl C.G., Takeuchi H., Nomoto K.;
 RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina
 fulica Ferussac containing a D-amino acid residue.";
 RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).
 RN (2)
 RP CHARACTERIZATION.
 RC STRAIN-PERUSSAC; TISSUE-Heart atrium;
 RX MEDLINE-91264856; Pubmed-1675568;
 RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,
 RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;
 RT "Purification of achatin-I from the atria of the African giant snail,
 Achatina fulica, and its possible function.";
 RL Biochem. Biophys. Res. Commun. 177:847-853(1991).
 RN (3)
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE-93014529; Pubmed-1399265;
 RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,
 RA Iwashita T., Nomoto K.;
 RT "Crystal structure and molecular conformation of achatin-I
 (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
 D-amino acid residue.";
 RL Int. J. Pept. Protein Res. 39:258-264(1992).
 CC -1- FUNCTION: NEUROEXCITATORY PEPTIDE: INCREASES THE IMPULSE FREQUENCY
 AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY
 NEURON (PON). ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE
 HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.
 CC NEURON (PON).
 CC CC
 DR PIR: A32480; A32480.
 KW Hormone; D-amino acid.
 FT MOD.RES 2
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 21.2%: Score 7; DB 1; Length 4;
 Best Local Similarity 50.0%: Pred. No. 1e-05; 1; Indels 0; Gaps 0;
 Matches 1; Conservative 1; Mismatches 0;

Oy 4 FS 5
 Db 2 FA 3

RESULT 12
 FMRF_MACNI STANDARD; PRT; 4 AA.
 ID FMRF_MACNI
 AC P01162;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
 OS Macrocallista nimbosa (Sun-ray clam),
 OS Neris virens (Sandworm),
 OS Hirudo medicinalis (Medicinal leech), and
 OS Helisoma trivolvis (Snail).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchla; Veneroida;
 OC Veneroidae; Veneridae; Macrocallista.
 OX NCB1_TaxID=6594, 6353, 6421, 27015;
 RN (1)
 RP SEQUENCE, AND SYNTHESIS.
 RC SPECIES-M. nimbosa; TISSUE-Cerebral pedal, and visceral ganglion;
 RX MEDLINE-77215956; Pubmed-877582;
 RA Price D.A., Greenberg M.J.;

```

RT "Structure of a molluscan cardioexcitatory neuropeptide."
RL Science 197;670-671(1977).
RN [2]
RN SEQUENCE, AND CHARACTERIZATION.
RC SPECIES=M.nimbosa; TISSUE=Ganglion;
RX MEDLINE=78012038; PubMed=909875;
RA Price D.A., Greenberg M.J.;
RT "Purification and characterization of a cardioexcitatory neuropeptide
from the central ganglia of a bivalve mollusc."
RL Prep. Biochem. 7;261-261(1977).
RN [3]
RN SEQUENCE.
RC SPECIES=N.virens; PubMed=2342992;
RX MEDLINE=90259866; PubMed=909875;
RA Krainiak K.G., Price D.A.;
RT "Authentic FMRFamide is present in the polychaete Nereis virens."
RL Peptides 11;75-77(1990).
RN [4]
RN SEQUENCE.
RC SPECIES=H.medicinalis; PubMed=1686933;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsenis M.A., Calabrese R.L.;
RT "Identification of Rramide neuropeptides in the medicinal leech."
RL Peptides 12;897-908(1991).
RN [5]
RN SEQUENCE.
RC SPECIES=H.trivoltis; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma
trivoltis."
RL Peptides 15;31-36(1994).
CC -1- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL
ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULARIZATION OF
CARDIAC CONTRACTION.
CC -1- SIMILARITY: BELONGS TO THE FAMP (FMRFAMIDE RELATED PEPTIDE)
FAMILY.
DR PIR: A01426; ECNK.
DR PIR: A60418; A60418.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 4;
Best Local Similarity 33.3%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 FSK 6
   1
   1
Db 1 FMR 3

RESULT 13
TUFF_HUMAN STANDARD: PRT; 4 AA.
AC P01858;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phagocytosis-stimulating peptide (Tuftsin).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE.
RX MEDLINE=72187087; PubMed=4112769;
RA Nishio K., Constantinopoulos A., Satoh P.S., Najjar V.A.;
RT "The characteristics, isolation and synthesis of the phagocytosis
stimulating peptide tuftsin."
RL Biochem. Biophys. Res. Commun. 47;172-179(1972).
RN [2]

```

```

RP IMMUNOGLOBULIN CLASS.
RX MEDLINE=68091045; PubMed=4169272;
RA Fidalgo B.V., Najjar V.A.;
RT "The physiological role of the lymphoid system. VI. The stimulatory
effect of leucophilic gamma globulin (leucokinin) on the phagocytic
activity of human polymorphonuclear leucocyte."
RL Biochemistry 6;3386-3392(1967).
CC -1- MISCELLANEOUS: AN IGG (CALLED LEUCOKININ) BINDS REVERSELY TO THE
CELL MEMBRANE OF NEUTROPHILS IN THE BLOOD. LEUCOKININASE ON THE
MEMBRANE RELEASES THE ACTIVE PEPTIDE TUFTSIN FROM THE GAMMA CHAIN.
CC TUFTSIN IS ESSENTIAL FOR MAXIMUM STIMULATION OF THE PHAGOCYTIC
ACTIVITY OF NEUTROPHILS.
DR PIR: A02147; A02147.
DR MIM: 191150; -
SQ SEQUENCE 4 AA; 501 MW; 74176321C0000000 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 4;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TR 2
   1
   1
Db 1 TK 2

RESULT 14
AL14_CARMA STANDARD: PRT; 5 AA.
AC P81817;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas (Common shore crab) (Green crab).
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RN SEQUENCE.
RP TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250;727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 5 5 AMIDATION (POTENTIAL).
SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 FS 5
   1
   1
Db 1 YS 2

RESULT 15
RE11_LITRU STANDARD: PRT; 5 AA.
ID RE11_LITRU
AC P82070;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidina l.1.
OS Litoria rubella (Desert tree frog).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-Skin secretion;
RA Steindorner S.T., Wabnitz P.A., Maugh R.J., Bowle J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT 'The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.';
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=598; METHOD=FAE.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

Query Match 21.2%; Score 7; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
I: 1;
DB 4 FA 5

Search completed: July 15, 2002, 13:38:10
Job time: 710 sec

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S19 BETA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC STRAIN-CV. ALMARO; TISSUE-LEAF;
RX MEDLINE=20435797; PubMed=10874039;
RA Yamaguchi K., von Knoblauch K., Sudramanian A.R.;
RT "The small ribosomal proteins. Identification of all the proteins in
the small subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 37:28455-28465(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD-ELECTROSPRAY.
CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
FORM IS THE MINOR BASIC FORM.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
CC InterPro: IPR002222; Ribosomal_S19.
DR Pfam: PF00203; Ribosomal_S19; PARTIAL.
DR PRINTS: PR00975; RIBOSOMALS19; PARTIAL.
DR PROSITE: PS00323; RIBOSOMAL_S19; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER
FT SEQUENCE 6 AA: 732 MW: 6333735A411C000 CRC64;

Query Match 39.4%; Score 13; DB 10; Length 6;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TRFSK 6
DB 1 TRSLRK 6

RESULT 3
ID 015903 PRELIMINARY; PRT; 7 AA.
AC 015903;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE (CLONE XP7E7B) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
arrayed cDNAs and cosmid libraries";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL: L32082; AAA73893.1; -.
FT NON_TER
FT SEQUENCE 7 AA: 849 MW: 6B040339CDD33DB0 CRC64;

Query Match 27.3%; Score 9; DB 4; Length 7;
Best Local Similarity 20.0%; Pred. No. 5.6e+05;

Matches 1; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 RIFSK 6
DB 2 KAFKR 6

RESULT 4
ID 095945 PRELIMINARY; PRT; 7 AA.
AC 095945;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INSIDE INTRON 5 (FRAGMENT).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=5254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
nucleotide sequence of the gene coding for subunit 1 of yeast
cytochrome oxidase";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL: Y00694; CA24066.1; -.
KW Mitochondrion.
FT NON_TER
FT SEQUENCE 7 AA: 859 MW: 75B7232362CDC460 CRC64;

Query Match 27.3%; Score 9; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 KL 7
DB 5 KL 6

RESULT 5
ID 09CSB3 PRELIMINARY; PRT; 7 AA.
AC 09CSB3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 0.7 KDA PROTEIN (FRAGMENT).
GN DIDI 10A-2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=21171025; PubMed=11277426;
RA Vercauteren I., Van Der Schueren E., Van Montagu M., Gheysen G.;
RT "Arabidopsis thaliana genes expressed in the early compatible
interaction with root-knot nematodes";
RL Mol. Plant Microbe Interact. 14:288-299(2001).
DR EMBL: AJ286350; CAB71014.2; -.
KW Hypothetical protein.
FT NON_TER
FT SEQUENCE 7 AA: 719 MW: 6732C7287EB325D0 CRC64;

Query Match 27.3%; Score 9; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 SK 6
DB 1 SK 2

RESULT 6
P82445 PRELIMINARY: PRT: 7 AA.
AC P82445;
DT 01-JUN-2000 (TREMblrel. 14, Created)
DT 01-JUN-2000 (TREMblrel. 14, Last sequence update)
DE 10 KDA CELL WALL PROTEIN (FRAGMENT).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. PETIT HAVANA;
RA Blee K.A., Bonham V.A., Mitchell G.P., Robertson D., Slabas A.R.,
RT Wojtaszek P., Bolwell G.P.;
RT "Proteomic study of secondary cell wall proteins from transformed
RT tobacco cultures";
RL Planta 010-0(2000).
CC -1- SUBCELLULAR LOCATION: CELL WALL.
CC -1- TISSUE SPECIFICITY: XYLEM.
KW Cell wall.
FT NON_TER
SQ SEQUENCE 7 AA: 758 MW; 69D2C1B62D1B2A0 CRC64;

Query Match 27.3%; Score 9; DB 10; Length 7;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 IF 4
DB 6 VF 7

RESULT 7
Q66205 PRELIMINARY: PRT: 7 AA.
AC Q66205;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE MEMBRANE PROTEIN (1 IS 3RD BASE IN CODON) (FRAGMENT).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FS772/70;
RX MEDLINE=86216185; Pubmed=2835592;
RA Britton P., Carmenes R.S., Page K.W., Garves D.J., Parra F.;
RT "Sequence of the Nucleoprotein Gene from a Virulent British Field
RT Isolate of Transmissible Gastroenteritis Virus and its Expression in
RT Saccharomyces Cerevisiae";
RL Mol. Microbiol. 2:89-99(1988).
DR EMBL: Y00542; CAA68606.1; -.
FT NON_TER
SQ SEQUENCE 7 AA: 869 MW; 72C691E727233B70 CRC64;

Query Match 27.3%; Score 9; DB 12; Length 7;

Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 KL 7
DB 2 KL 3

RESULT 8
O99182 PRELIMINARY: PRT: 7 AA.
AC O99182;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CYTOCHROME OXIDASE I (FRAGMENT).
GN COI.
OS Gnatholebias zonatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.
OX NCBI_TaxID=135316;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20072928; Pubmed=10603257;
RA Murphy W.J., Thomerson J.E., Collier G.E.;
RT "Phylogeny of the Neotropical killifish family Rivulidae
RT (Cyprinodontiformes, Aplocheiloidae) inferred from mitochondrial DNA
RT sequences";
RL Mol. Phylogenet. Evol. 13:289-301(1999).
DR EMBL: AF002591; AAD01074.1; -.
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 7 AA: 899 MW; 672721F6C8572030 CRC64;

Query Match 24.2%; Score 8; DB 8; Length 7;
Best Local Similarity 20.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 IFSKL 7
DB 2 LYQHL 6

RESULT 9
O98866 PRELIMINARY: PRT: 7 AA.
AC O98866;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
DE CYTOCHROME B/F SUBUNIT IV (FRAGMENT).
OS Spinacia oleracea (Spinach).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86120353; Pubmed=3003688;
RA Siben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT protein S11 and RNA polymerase alpha-subunit";
RL Nucleic Acids Res. 14:1029-1044(1986).
DR EMBL: X03496; CAA27215.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA: 907 MW; 644729D77409CA20 CRC64;

Query Match 24.2%; Score 8; DB 8; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RIF 4
1
1
DB 3 RPF 5

RESULT 10

ID P82070 PRELIMINARY; PRT; 5 AA.
AC P82070; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
RT Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=598; METHOD=FAB.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

Query Match 21.2%; Score 7; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
1
1
DB 4 FA 5

RESULT 11

ID P82071 PRELIMINARY; PRT; 5 AA.
AC P82071; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
RT Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.

CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=626; METHOD=FAB.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

Query Match 21.2%; Score 7; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
1
1
DB 4 FA 5

RESULT 12

ID P82072 PRELIMINARY; PRT; 5 AA.
AC P82072; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE-SKIN SECRETION;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella', the skin peptide profile as a probe for the study
of evolutionary trends of amphibians.";
RT Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=655; METHOD=FAB.
KW Amphibian skin; Amidation.
FT MOD.RES 5
SQ SEQUENCE 5 AA; 655 MW; 71A9C9C810300000 CRC64;

Query Match 21.2%; Score 7; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 FS 5
1
1
DB 4 FT 5

RESULT 13

ID P82073 PRELIMINARY; PRT; 5 AA.
AC P82073; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE-SKIN SECRETION;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog

```

RT Litoria electrica. Comparison with the skin peptides from Litoria
RL rubella.".
CC Aust. J. Chem. 52:0-0(1999).
CC -I- FUNCTION: CARIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTI-BIOTIC ACTIVITY.
CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KM Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C62A00000 CRC64;

Query Match 21.2%; Score 7; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 FS 5
   1:
DB 4 FT 5

RESULT 14
P82181 PRELIMINARY; PRT; 6 AA.
ID P82181;
AC P82181;
DT 01-JUN-2000 (TREMBLREL. 14, Created)
DT 01-JUN-2000 (TREMBLREL. 14, last sequence update)
DT 01-JUN-2001 (TREMBLREL. 17, last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_Taxid=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALMARO. TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -I- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -I- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -I- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -I- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KM Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 21.2%; Score 7; DB 10; Length 6;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TR 2
   1:
DB 5 TK 6

RESULT 15
P82182 PRELIMINARY; PRT; 6 AA.
ID P82182;
AC P82182;
DT 01-JUN-2000 (TREMBLREL. 14, Created)
DT 01-JUN-2000 (TREMBLREL. 14, last sequence update)
DT 01-JUN-2001 (TREMBLREL. 17, last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 GAMMA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_Taxid=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALMARO. TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -I- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -I- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -I- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -I- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; Ribosomal_L10.
DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KM Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 21.2%; Score 7; DB 10; Length 6;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TR 2
   1:
DB 5 TK 6

```

Search completed: July 15, 2002, 13:37:39
Job time: 729 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:25:24 ; Search time 107.71 Seconds
(without alignments)
7.219 Million cell updates/sec

Title: US-09-712-819A-2
Perfect score: 33
Sequence: 1 TRFSKL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 52936

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	23	69.7	7	20	AAW92345
2	20	60.6	7	16	AAW87299
3	20	60.6	7	21	AAW28929
4	19	57.6	7	18	AAW18611
5	18	54.5	4	21	AAW13882
6	18	54.5	6	17	AAW08723
7	18	54.5	6	17	AAW08724
8	18	54.5	6	19	AAW68419
9	18	54.5	6	21	AAW13880
10	18	54.5	6	21	AAW13881
11	18	54.5	7	22	AAW98766

12	17	51.5	5	18	AAW18503	Amino-terminal pep
13	17	51.5	6	19	AAW68421	Peptide #3 from cy
14	17	51.5	6	20	AAW68428	Peptide #10 from c
15	17	51.5	6	20	AAW17746	Calreticulin bindi
16	17	51.5	7	13	AAW24501	Monomeric repeatin
17	17	51.5	7	20	AAW40797	Amino acid loop 13
18	17	51.5	7	21	AAW30137	Scaffold protein S
19	17	51.5	7	21	AAW22771	FAUP receptor-targ
20	16	48.5	4	5	AAW40463	N-terminal polynhyd
21	16	48.5	4	17	AAW67776	Tetrapeptide usetu
22	16	48.5	5	16	AAW50125	Bovine, ovine and
23	16	48.5	5	16	AAW78975	Peptide inhibitor
24	16	48.5	5	18	AAW12483	Interleukin-6 anta
25	16	48.5	5	21	AAW06421	Claudin-1 cell adh
26	16	48.5	6	15	AAW50143	Ovine growth hormo
27	16	48.5	6	15	AAW50144	Ovine growth hormo
28	16	48.5	6	16	AAW89186	3' end of alkaline
29	16	48.5	6	19	AAW75336	Hexapeptide #1 bin
30	16	48.5	6	19	AAW75268	Enzyme inhibitor P
31	16	48.5	6	21	AAW56849	Claudin-1 cell adh
32	16	48.5	6	21	AAW06422	Claudin-1 cell adh
33	16	48.5	6	21	AAW06424	Claudin-1 cyclic c
34	16	48.5	6	21	AAW06510	Peptide used to de
35	16	48.5	6	21	AAW69210	Peptide derived fr
36	16	48.5	6	21	AAW68614	Breast cancer susc
37	16	48.5	7	15	AAW50149	Ovine growth hormo
38	16	48.5	7	15	AAW50150	Ovine growth hormo
39	16	48.5	7	15	AAW50126	Bovine, ovine and
40	16	48.5	7	16	AAW77517	Cytochrome-P450-1p
41	16	48.5	7	20	AAW40748	S4 derivative #22,
42	16	48.5	7	20	AAW33959	Dynorphin A(6-12)
43	16	48.5	7	21	AAW30086	Scaffold protein S
44	16	48.5	7	21	AAW06423	Claudin-1 cell adh
45	16	48.5	7	21	AAW06423	Claudin-1 cell adh

ALIGNMENTS

RESULT 1	
AAW92345	AAW92345 standard; peptide: 7 AA.
ID	AAW92345
AC	AAW92345;
DT	09-APR-1999 (first entry).
DE	Latent fluorescent peptide #9.
KW	Fluorescent peptide; masking group; fluorescent group; sensitive; stable;
KW	phosphatase activity; phosphoric acid.
OS	Synthetic.
PN	JPL1012297-A.
PD	19-JAN-1999.
PF	21-APR-1998; 98JP-0126684.
PR	30-APR-1997; 97JP-0126463.
PA	(NISHU) NISHIKATA S.
DR	WPI, 1999-148570/13.
PT	New latent fluorescent peptide(s) - useful for determining protein
PT	phosphatase activity
PS	Claim 5, Page 9; 11pp; Japanese.
CC	AAW92337-992350 are novel latent fluorescent (optionally protected)
CC	peptides having a self fluorescent group and masking group in the

CC molecule and phosphoric acid groups in the masking group and between the
 CC fluorescent and masking groups. The invention also describes peptide
 CC derivatives of formula [A1-X-A2], especially composed of 3-42 amino
 CC acids; X = amino acid residue comprising phosphoric acid residue; A1,
 CC A2 = (optionally protected) amino acid residue, (optionally protected)
 CC peptide chain, one comprises a masking group at the terminal or side
 CC chain and the other has a self fluorescent group capable of being masked
 CC with the intramolecular masking group. The peptides are useful for
 CC determining protein phosphatase activity by the changes of intensity of
 CC fluorescence. The fluorescent peptides are stable reagents for sensitive
 CC and rapid determination of protein phosphatase activity in crude sample
 CC even contaminated sample containing phosphoric acid ion.

CC Sequence 7 AA;

Query Match 69.7%; Score 23; DB 20; Length 7;
 Best Local Similarity 80.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIFSK 6
 Db 2 rvfsk 6

RESULT 2
 AAR87299
 ID AAR87299 standard; peptide: 7 AA.

AC AAR87299;

DT 16-MAR-1996 (first entry)

DE Plasmodium falciparum HLA-B17 epitope 1s55.

KM Liver stage antigen: LSA-1; human leucocyte antigen: HLA; class I;

KW HLA-B17; epitope: malaria; vaccine; CTL induction;

KV cytotoxic T lymphocyte.

OS Plasmodium falciparum.

PN WO9526982-A2.

PD 12-OCT-1995.

PF 31-MAR-1995; 95MO-GB00737.

PR 31-MAR-1994; 94GB-0006492.

PA (ISIS-) ISIS INNOVATION LTD.

PI Aldo M, Allsopp CEM, Hill AVS, Lalvani A, Plebanski M;

PI Whittle HC;

DR WPI; 1995-358584/46.

PT Plasmodium falciparum peptide(s) - useful in vaccine compositions
 PT for immunising against malaria

PS Claim 1; Page 19; 23pp; English.

CC Cytotoxic T lymphocytes from malaria-exposed Gambian individuals
 CC with HLA-B17 showed significant lysis of a large pool of peptides
 CC (AAR87287-R87299) derived from Plasmodium falciparum liver stage
 CC antigen-1. These peptides will be useful in a malaria vaccine.

CC Sequence 7 AA;

Query Match 60.6%; Score 20; DB 16; Length 7;
 Best Local Similarity 57.1%; Pred. No. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 Db 1 tkyfml 7

RESULT 3
 AAB28929
 ID AAB28929 standard; peptide: 7 AA.

AC AAB28929;

DT 25-JAN-2001 (first entry)

DE Geminivirus Rep peptide sequence #31.

KM Geminivirus; replication associated protein; rep; iteron;

KW antagonists; plant.

OS Geminivirus.

PN WO200043494-A2.

PD 27-JUL-2000.

PF 27-JAN-2000; 2000WO-US01849.

PR 26-JAN-1999; 99US-0117285.

PA (SCRI) SCRIPPS RES INST.

PI Faquet C, Chatterji A;

DR WPI; 2000-499224/44.

PT Producing plants resistant to geminivirus, and inhibiting geminivirus
 PT replication in plants, by introducing replication associated protein
 PT iteron antagonists into the plant, plant cells or propagules -

PS Claim 5; Page 106; 172pp; English.

CC The present invention relates to methods for producing plants resistant
 CC to geminivirus, involving introducing a geminivirus replication
 CC associated protein (Rep)-iteron antagonist into a plant. The antagonist
 CC is a nucleotide sequence of a geminivirus iteron capable of binding to
 CC a Rep protein or a defective Rep which has a conserved geminivirus
 CC iteron binding site. The present sequence is a geminivirus Rep peptide
 CC sequence.

CC Sequence 7 AA;

Query Match 60.6%; Score 20; DB 21; Length 7;
 Best Local Similarity 60.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIFSK 6
 Db 2 rvysk 6

RESULT 4
 AAW18611
 ID AAW18611 standard; peptide: 7 AA.

AC AAW18611;

DT 04-MAR-1998 (first entry)

DE Aged band 3 peptide (residues 534-540) epitope.

KW Band 3 protein; antibody; aging antigenic site; Alzheimer's disease;

KX phosphorylation; detection; epitope.

```

OS Homo sapiens.
XX
XX WO9726537-A1.
XX
XX 24-JUL-1997.
XX
XX 13-DEC-1996; 96WO-US20465.
XX
XX 19-JAN-1996; 96US-0010250.
XX
XX (RESE ) RESEARCH CORP TECHNOLOGIES INC.
XX
XX Kay MMB;
XX
XX WPI: 1997-385478/35.
XX
XX Detecting Alzheimer's disease using antibody that recognises aged
XX band 3 protein in tissues - or from reduced degree of band 3 protein
XX phosphorylation, can be applied to blood or brain samples
XX
XX Disclosure; Page 12; 45pp; English.
XX
XX This is an aging antigenic band 3 peptide (residues 534-540) to which a
XX specific antibody can bind to. Band 3 is a ubiquitous anion-exchange
XX protein and ages as cells and tissues age. Antibodies have been developed
XX against this aged band 3. These antibodies bind to distinct regions of
XX band 3 in old cells (aging antigenic sites) but not middle aged or young
XX cells. This can be used for detecting Alzheimer's disease. A tissue
XX sample containing band 3 from a patient suspected of having Alzheimer's
XX disease is treated with an antibody that can differentiate between the
XX Alzheimer's (aged) and normal band 3, under complex-forming conditions
XX and detecting any complex formed. A tissue containing band 3 from a
XX healthy control is treated in a similar manner and the amounts of complex
XX formed are compared. A significantly greater formation of complex in the
XX suspect sample as compared with that of the control is indicative of
XX Alzheimer's disease. The disease can also be detected by comparing the
XX degrees of phosphorylation of band 3 or its degradation products in
XX suspect and control samples. A significant decrease in phosphorylation
XX in the suspect sample indicates Alzheimer's disease.
XX
XX Sequence 7 AA:
SQ
Query Match 57.6%; Score 19; DB 18; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 FSKL 7
DB 4 fskl 7

```

```

XX 24-FEB-1999; 99US-0121327.
XX
XX 24-FEB-1999; 99US-0256970.
XX
XX 23-SEP-1999; 99US-0155492.
XX
XX 23-SEP-1999; 99US-0404431.
XX
XX 23-FEB-2000; 2000US-0511956.
XX
XX (ACOR-) ACORDA THERAPEUTICS.
XX
XX Schachner M, Neuberger TJ, Herzberg U, Simon M;
XX
XX WPI: 2000-579163/54.
XX
XX New isolated peptide which mimics a carbohydrate epitope is useful for
XX neuroprotection
XX
XX Claim 9; Page 124; 214pp; English.
XX
XX The present invention relates to L2/HNK1 carbohydrate epitope mimic
XX peptides. The L2/HNK1 epitope is predominantly expressed on glycolipids
XX and glycoproteins from nervous tissue. Peptides that mimic the L2/HNK1
XX epitope were isolated by screening phage peptide display libraries with
XX antibodies to L2/HNK1. The present sequence is the homologous region
XX between consensus sequences of mimic peptides isolated by different
XX antibodies. The peptides may be used for enhancing memory and for
XX treating Alzheimer's disease or dementia. They may also be used for
XX treating apoptosis, necrosis, Parkinson's disease, multiple sclerosis
XX and acute and chronic spinal cord injury.
XX
XX Sequence 4 AA:
SQ
Query Match 54.5%; Score 18; DB 21; Length 4;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRIF 4
DB 1 trlf 4

```

```

RESULT 6
AAW08723
ID AAW08723 standard; peptide; 6 AA.
XX
XX AAW08723;
AC
XX 28-FEB-1997 (first entry)
DT
XX
XX Trypsin inhibitory peptide #4.
DE
XX
XX Ligand: synthetic combinatorial peptide library; hexamer; antibody;
XX antigen; receptor; inhibitor; trypsin.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 1
FT Modified-site /label= OTHER
FT /note= "Ac-Met"
FT Modified-site /label= OTHER
FT /note= "Thr-NH2"
FT
FT
XX US5556762-A.
XX
XX 17-SEP-1996.
PD
XX
XX 21-NOV-1990; 90US-0617023.
PF
XX
XX 11-SEP-1992; 92US-0943709.
PR
XX 21-NOV-1990; 90US-0617023.
PR
XX 16-MAY-1991; 91US-0701658.
PR

```

PR 19-NOV-1991; 91US-0797551.
XX
XX (HOUG-) HOUGHTEN PHARM INC.
XX
XX
PI Appel JR, Houghten RA, Pinilla C;
XX
XX WPI: 1996-432985/43.
DR
XX
XX Identifying oligopeptide ligands for an acceptor - by scanning
PT synthetic peptide combinatorial libraries comprising
PT self-solubilizing, unsupported mixed oligopeptide(s)
XX
XX
PS Example 8; Column 59; 75pp; English.
XX
XX The invention relates to a method of identifying oligopeptide ligands to
CC a protein by scanning synthetic combinatorial peptide libraries (SCPL).
CC The SCPL comprise sets of hexamers which contain either one of 6
CC predetermined amino acids at one predetermined position in the hexamer
CC and each set may have one predetermined amino acid at 1 of 6
CC predetermined pos. in the hexamer. The method is useful for identifying
CC biologically active sequences of e.g. pharmaceutical use. The peptides
CC esp. inhibit the binding of e.g. an antibody to its antigen, a ligand to
CC its receptor, etc.
CC The peptides AAW08720-25 are N-terminally acylated and C-terminally
CC amidated peptides, isolated from a SCPL, which inhibit trypsin.
CC This peptide inhibited trypsin with an IC50 of 133 micromole.
XX
XX
SQ Sequence 6 AA:

Query Match 54.5%; Score 18; DB 17; Length 6;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFS 5
1:11:
Db 2 tkift 6

RESULT 7
AAW08724
ID AAW08724 standard; peptide; 6 AA.
XX
XX AAW08724;
AC
XX
XX 28-FEB-1997 (first entry)
DT
XX
XX Trypsin inhibitory peptide #5.
DE
XX
XX Ligand; synthetic combinatorial peptide library; hexamer; antibody;
KW antigen; receptor; inhibitor; trypsin.
XX
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= OTHER
FT /note= "Ac-Thr"
FT Modified-site 6
FT /label= OTHER
FT /note= "Thr-NH2"
XX
XX
XX US5556762-A.
XX
XX
XX 17-SEP-1996.
PD
XX
XX 21-NOV-1990; 90US-0617023.
PF
XX
XX 11-SEP-1992; 92US-0943709.
PR 21-NOV-1990; 90US-0617023.
PR 16-MAY-1991; 91US-0701658.
PR 19-NOV-1991; 91US-0797551.
XX
XX
PT

PA (HOUG-) HOUGHTEN PHARM INC.
XX
XX
XX Appel JR, Houghten RA, Pinilla C;
XX
XX
DR WPI: 1996-432985/43.
XX
XX
XX Identifying oligopeptide ligands for an acceptor - by scanning
PT synthetic peptide combinatorial libraries comprising
PT self-solubilizing, unsupported mixed oligopeptide(s)
XX
XX
PS Example 8; Column 59; 75pp; English.
XX
XX The invention relates to a method of identifying oligopeptide ligands to
CC a protein by scanning synthetic combinatorial peptide libraries (SCPL).
CC The SCPL comprise sets of hexamers which contain either one of 6
CC predetermined amino acids at one predetermined position in the hexamer
CC and each set may have one predetermined amino acid at 1 of 6
CC predetermined pos. in the hexamer. The method is useful for identifying
CC biologically active sequences of e.g. pharmaceutical use. The peptides
CC esp. inhibit the binding of e.g. an antibody to its antigen, a ligand to
CC its receptor, etc.
CC The peptides AAW08720-25 are N-terminally acylated and C-terminally
CC amidated peptides, isolated from a SCPL, which inhibit trypsin.
CC This peptide inhibited trypsin with an IC50 of 184 micromole.
XX
XX
SQ Sequence 6 AA:

Query Match 54.5%; Score 18; DB 17; Length 6;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFS 5
1:11:
Db 2 tkift 6

RESULT 8
AAW68419
ID AAW68419 standard; peptide; 6 AA.
XX
XX AAW68419;
AC
XX
XX 23-NOV-1998 (first entry)
DT
XX
XX Peptide #1 from cyclic screening method for physiological activity.
DE
XX
XX Screening method; physiological activity; hexapeptide; algorithm;
KW crossover; mutation; enzyme inhibitor; drug; foodstuff.
XX
XX
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "acylated N-terminus"
FT Modified-site 6
FT /note= "amidated C-terminus"
PN W09804580-A1.
XX
XX
XX 05-FEB-1998.
PD
XX
XX 23-JUL-1997; 97WO-JP02535.
XX
XX
XX 26-JUL-1996; 96JP-0198096.
PR
XX
XX (YAMA) YAMANOUCHI PHARM CO LTD.
PA
XX
XX karube M, yokabayashi Y;
PI
XX
XX WPI: 1998-130616/12.
DR
XX
XX
PT Method for screening peptide(s) with high physiological activity -

PT by determining activity, dropping least active, changing sequence in
PT remaining peptide(s) and cycling through this sequence to obtain
PT peptide(s) with high activity
XX
PS Disclosure; Page 9; 19pp; Japanese.
XX
CC Peptides AAM6819-W68434 represent peptides obtained by the cyclic
CC screening method of the invention. The method comprises determining
CC the level of a desired physiological activity in a series of
CC hexapeptides and selecting those with the highest activities. The
CC peptides are then processed using a genetic algorithm computer program
CC which: (i) exchanges one or more amino acid residues between pairs of
CC peptides (crossover) and/or (ii) substitutes amino acid residues in a
CC peptide for different ones (mutation) (preferably with a mutation
CC frequency of about 3% of the total number of residues). The altered
CC peptides are synthesised and the screening cycle is repeated until
CC peptides with the desired level of activity are generated. Peptides with
CC a high activity (e.g. as enzyme inhibitors) can be used to produce
CC compositions for use in the drug, foodstuff and other industries.
XX
SQ Sequence 6 AA;

Query Match 54.5%; Score 18; DB 19; Length 6;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIPS 5
|:|:|
Db 2 tkltf 6

RESULT 9
AAB13880
ID AAB13880 standard; peptide; 6 AA.
XX
AC AAB13880;
XX
DT 07-NOV-2000 (first entry)
XX
DE Consensus sequence for mimic peptides selected by HNKL antibody.
XX
XX L2/HNKL; epitope; nervous tissue; apoptosis; necrosis; antibody;
KW Parkinson's disease; multiple sclerosis; spinal cord injury;
KW Alzheimer's disease; dementia.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 6 /label= Val, Phe
XX
XX NO200050447-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000MO-US04730.
XX
XX
PR 24-FEB-1999; 99US-0121327.
PR 24-FEB-1999; 99US-0256970.
PR 23-SEP-1999; 99US-0155492.
PR 23-SEP-1999; 99US-0404431.
PR 23-FEB-2000; 2000US-0511956.
XX
XX (ACOR-) ACORDA THERAPEUTICS.
XX
PI Schachner M, Neuberger TJ, Herzberg U, Simon M;
XX
XX WPI; 2000-579163/54.
XX
XX New isolated peptide which mimics a carbohydrate epitope is useful for
PT neuroprotection -
XX

PS Claim 7; Page 124; 214pp; English.
XX
XX The present invention relates to L2/HNKL carbohydrate epitope mimic
CC peptides. The L2/HNKL epitope is predominantly expressed on glycolipids
CC and glycoproteins from nervous tissue. Peptides that mimic the L2/HNKL
CC epitope were isolated by screening phage peptide display libraries with
CC antibodies to L2/HNKL. The present sequence is the consensus sequence
CC for mimic peptides selected by HNKL antibody. The peptides may be used
CC for enhancing memory and for treating Alzheimer's disease or dementia.
CC They may also be used for treating apoptosis, necrosis, Parkinson's
CC disease, multiple sclerosis and acute and chronic spinal cord injury.
XX
SQ Sequence 6 AA;

Query Match 54.5%; Score 18; DB 21; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIF 4
|:|:|
Db 1 trlf 4

RESULT 10
AAB13881
ID AAB13881 standard; peptide; 6 AA.
XX
AC AAB13881;
XX
DT 07-NOV-2000 (first entry)
XX
DE L2/HNKL epitope mimic peptide homology sequence #1.
XX
XX
KW L2/HNKL; epitope; nervous tissue; apoptosis; necrosis; antibody;
KW Parkinson's disease; multiple sclerosis; spinal cord injury;
KW Alzheimer's disease; dementia.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 5 /note="Arg or nothing"
XX
XX NO200050447-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000MO-US04730.
XX
XX
PR 24-FEB-1999; 99US-0121327.
PR 24-FEB-1999; 99US-0256970.
PR 23-SEP-1999; 99US-0155492.
PR 23-SEP-1999; 99US-0404431.
PR 23-FEB-2000; 2000US-0511956.
XX
XX (ACOR-) ACORDA THERAPEUTICS.
XX
PI Schachner M, Neuberger TJ, Herzberg U, Simon M;
XX
XX WPI; 2000-579163/54.
XX
XX New isolated peptide which mimics a carbohydrate epitope is useful for
PT neuroprotection -
XX
XX Claim 8; Page 124; 214pp; English.
XX
XX The present invention relates to L2/HNKL carbohydrate epitope mimic
CC peptides. The L2/HNKL epitope is predominantly expressed on glycolipids
CC and glycoproteins from nervous tissue. Peptides that mimic the L2/HNKL
CC epitope were isolated by screening phage peptide display libraries with
CC antibodies to L2/HNKL. The present sequence is the homologous region
CC between consensus sequences of mimic peptides isolated by different

CC antibodies. The peptides may be used for enhancing memory and for
 CC treating Alzheimer's disease or dementia. They may also be used for
 CC treating apoptosis, necrosis, Parkinson's disease, multiple sclerosis
 CC and acute and chronic spinal cord injury.
 XX

SO Sequence 6 AA;

Query Match 54.5%; Score 18; DB 21; Length 6;
 Best Local Similarity 75.0%; Pred. No. 6.4e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIF 4
 11:1
 Db 1 trlf 4

RESULT 11

ID AAG98766

AC AAG98766 standard; peptide; 7 AA.

AC AAG98766;

DT 21-SEP-2001 (first entry)

XX Human cell death protective cDNA clone CNI-00721 ORF19 peptide, SEQ:330.

XX Cell death protective; apoptosis; necrosis; human; drug screening;

KW cell death-associated disorder; central nervous system disorder;

KW psychiatric disorder; neurological disorder; ischemia-related disorder;

KW stroke; cerebral infarction; ischemic encephalopathy;

KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;

KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;

KW vascular disease; ophthalmological disorder; diabetic retinopathy;

KW macular degeneration; hypertension; myocardial infarction;

KW atherosclerosis; respiratory disorder; asthma; transgenic animal;

KW chronic obstructive pulmonary disease; neoplastic condition; cancer;

KW benign tumour; anaemia; gastrointestinal disorder; gastritis;

KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;

KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;

KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;

KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS.

XX Homo sapiens.

OS WO200145638-A2.

PN 28-JUN-2001.

XX 11-DEC-2000; 2000WO-US33547.

PR 14-DEC-1999; 99US-0461697.

XX (COGE-) COGENT NEUROSCIENCE INC.

XX Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;

PI WPI; 2001-390297/41.

DR N-PSDB; AAH84281, AAH84300.

XX Novel protective sequence polynucleotides and polypeptides, used to

PT identify modulators of their expression and activity, which are used in

PT to treat central nervous system conditions, diseases and disorders -

PS Claim 1: Flg 11S; 325pp; English.

XX Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
 CC protect against cell death (i.e., apoptosis or necrosis). Sequences
 CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
 CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
 CC while the remaining nucleic acid sequences within the range given above
 CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
 CC AAG98610-AAG98829 represent the polypeptides encoded by the cell death

CC protective ORFs. The cell death protective cDNA clones are able to
 CC prevent, delay or reverse progression through the apoptotic or necrotic
 CC pathways when injected into a cell predisposed to or undergoing cell
 CC death. The cell death protective nucleic acids and polypeptides can be
 CC used in the diagnosis and treatment of disorders associated with cell
 CC death, and to screen for compounds which modulate their activity or
 CC expression. Such modulators, preferably a small organic molecule, an
 CC antibody, a ribozyme, or an antisense molecule, can also be used to treat
 CC cell death-related diseases. Such diseases include those associated with
 CC the central nervous system including psychiatric or neurological
 CC disorders, especially ischemia-related conditions such as strokes, and
 CC also includes neurodegenerative disorders such as Alzheimer's disease,
 CC Huntington's disease, or Parkinson's disease. The modulators may also be
 CC used to treat infections such as meningitis, malaria, or trypanosomiasis;
 CC vascular diseases such as ischemic encephalopathy or cerebral
 CC infarction; eye conditions such as diabetic retinopathy or macular
 CC degeneration; hypertension; myocardial infarction; atherosclerosis;
 CC respiratory conditions such as asthma or chronic obstructive pulmonary
 CC disease; neoplastic conditions such as cancers or benign tumours; blood
 CC cell conditions such as anaemia; gastrointestinal conditions such as
 CC gastritis or ulcerative colitis; liver conditions such as biliary
 CC cirrhosis; kidney disorders such as glomerulonephritis; cystitis;
 CC endometriosis; endocrine disorders such as Grave's disease or Hashimoto's
 CC thyroiditis; skin conditions such as dermatitis or urticaria; or immune
 CC system disorders such as acquired immunodeficiency syndrome (AIDS). The
 CC nucleic acids may additionally be used to generate animal models of
 CC cell death-associated disorders. The present sequence represents a
 CC cell death protective polypeptide.

SO Sequence 7 AA;

Query Match 54.5%; Score 18; DB 22; Length 7;
 Best Local Similarity 60.0%; Pred. No. 6.4e+05;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 IFSKL 7
 11:1
 Db 2 ifskl 6

RESULT 12

ID AAM18503

AC AAM18503 standard; peptide; 5 AA.

AC AAM18503;

DT 19-FEB-1998 (first entry)

XX Amino-terminal peptide 5 associated with novel helicase.

KW Nucleic acid binding protein; helicase; leflunomide; assaying;
 KW anticancer; anticholesterolic; immunosuppressant; sequencing;
 KW antiinflammatory; antiviral; antifungal; antibacterial; treatment;

KW Alzheimer's disease; cancer; rheumatism; arthritis; determination;

KW atherosclerosis; osteoporosis; acute infection; chronic infection;

KW autoimmune disease; diabetes; organ transplant; isolation;
 KW amino-terminal.

XX Homo sapiens.

OS DE19545126-A1.

PN 05-JUN-1997.

XX 04-DEC-1995; 95DE-1045126.

PR 04-DEC-1995; 95DE-1045126.

XX (FARR) HOECHST AG.

PA Bartlett R, Kirschbaum B, Mueller S;

PI

XX WPI: 1997-299388/28.
DR
XX New nucleic acid binding protein with helicase activity - is
PT strongly induced by leflunomide, used to isolate specific binding
PT RNA and for identifying substances with anticancer, antiviral etc.
PT activities
XX
XX Example 4; Page 15; 28pp; German.
XX
XX The present sequence is an amino-terminal peptide associated with a
CC novel nucleic acid binding protein with helicase activity, the mRNA
CC of which (or its translation products) is strongly expressed in
CC presence of leflunomide, i.e.
CC 4-(N-(4-trifluoromethylphenyl)-5-methyl-isoazole, or a compound
CC with similar activity. The helicase can be used in assay systems
CC to identify/discover anticancer, antiatherosclerotic,
CC immunosuppressing, antiinflammatory, antiviral, antifungal and
CC antibacterial agents, e.g. to treat Alzheimer's disease, cancer,
CC rheumatism, arthritis, atherosclerosis, osteoporosis,
CC acute/chronic infections, autoimmune disease, diabetes and
CC complications of organ transplants, and to isolate or determine the
CC sequences of specific binding RNA.
CC
XX Sequence 5 AA:
SQ

Query Match 51.5%; Score 17; DB 18; Length 5;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RIFSK 6
:|||
Db 1 klfqk 5

RESULT 13
AAW68421
ID AAW68421 standard; peptide; 6 AA.
XX
AC AAW68421;
XX
DT 23-NOV-1998 (first entry)
XX
XX Peptide #3 from cyclic screening method for physiological activity.
DE
XX Screening method; physiological activity; hexapeptide; algorithm;
KM crossover; mutation; enzyme inhibitor; drug; foodstuff.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "acylated N-terminus"
FT Modified-site 6 /note= "amidated C-terminus"
FT
FT
XX W09804580-A1.
XX
XX 05-FEB-1998.
XX
XX 23-JUL-1997; 97WO-JP02535.
XX
XX 26-JUL-1996; 96JP-0198096.
XX
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX Karube M, Yokobayashi Y;
XX WPI: 1998-130616/12.
XX
XX Method for screening peptide(s) with high physiological activity -
PT by determining activity, dropping least active, changing sequence in
PT

PT remaining peptide(s) and cycling through this sequence to obtain
PT peptide(s) with high activity
XX
XX
PS Disclosure; Page 9; 19pp; Japanese.
XX
XX Peptides AAW68419-W68434 represent peptides obtained by the cyclic
CC screening method of the invention. The method comprises determining
CC the level of a desired physiological activity in a series of
CC hexapeptides and selecting those with the highest activities. The
CC peptides are then processed using a genetic algorithm computer program
CC which (i) exchanges one or more amino acid residues between pairs of
CC peptides (crossover) and/or (ii) substitutes amino acid residues in a
CC peptide for different ones (mutation) (preferably with a mutation
CC frequency of about 3% of the total number of residues). The altered
CC peptides are synthesized and the screening cycle is repeated until
CC peptides with the desired level of activity are generated. Peptides with
CC a high activity (e.g. as enzyme inhibitors) can be used to produce
CC compositions for use in the drug, foodstuff and other industries.
XX
XX Sequence 6 AA:
SQ

Query Match 51.5%; Score 17; DB 19; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIF 4
:|||
Db 1 tklf 4

RESULT 14
AAW68428
ID AAW68428 standard; peptide; 6 AA.
XX
AC AAW68428;
XX
DT 23-NOV-1998 (first entry)
XX
XX Peptide #10 from cyclic screening method for physiological activity.
DE
XX Screening method; physiological activity; hexapeptide; algorithm;
KM crossover; mutation; enzyme inhibitor; drug; foodstuff.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "acylated N-terminus"
FT Modified-site 6 /note= "amidated C-terminus"
FT
FT
XX W09804580-A1.
XX
XX 05-FEB-1998.
XX
XX 23-JUL-1997; 97WO-JP02535.
XX
XX 26-JUL-1996; 96JP-0198096.
XX
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX Karube M, Yokobayashi Y;
XX WPI: 1998-130616/12.
XX
XX Method for screening peptide(s) with high physiological activity -
PT by determining activity, dropping least active, changing sequence in
PT remaining peptide(s) and cycling through this sequence to obtain
PT peptide(s) with high activity
XX
XX Disclosure; Page 9; 19pp; Japanese.
XX

CC Peptides AAM68419-H68414 represent peptides obtained by the cyclic
CC screening method of the invention. The method comprises determining
CC the level of a desired physiological activity in a series of
CC hexapeptides and selecting those with the highest activities. The
CC peptides are then processed using a genetic algorithm computer program
CC which: (i) exchanges one or more amino acid residues between pairs of
CC peptides (crossover) and/or (ii) substitutes amino acid residues in a
CC peptide for different ones (mutation) (preferably with a mutation
CC frequency of about 3% of the total number of residues). The altered
CC peptides are synthesised and the screening cycle is repeated until
CC peptides with the desired level of activity are generated. Peptides with
CC a high activity (e.g. as enzyme inhibitors) can be used to produce
CC compositions for use in the drug, foodstuff and other industries.

XX
SQ Sequence 6 AA:

Query Match 51.5%; Score 17; DB 19; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRIF 4
Db 2 kRf 5

RESULT 15

AA17746
ID AA17746 standard; peptide: 6 AA.

AC AA17746;

DT 10-AUG-1999 (first entry)

DE Calreticulin binding peptide SEQ ID NO:65.

KW Calreticulin; binding; hormone responsiveness; modulation; mimetic;
KW inhibitor; gene therapy; cancer; osteoporosis; pharmaceutical;

KW chronic inflammatory disease; promyelocytic leukaemia; arthritis;
KW bone disease; hormone receptor binding.

OS Synthetic.

PN WO9905172-A2.

PD 04-FEB-1999.

PF 24-JUL-1998; 98WO-CA00715.

PR 24-JUL-1997; 97US-0900241.

PA (DEDH/) DEDHAR S.
PA (DOER/) DOERSEN C W.

PA (MAZU/) MAZUR A W.

PI Dedhar S, Doersen CW, Mazur AW;

DR WPI: 1999-142854/12.

PT New peptides that bind calreticulin and modulate gene expression -
PT are activated by hormone receptors, useful in the treatment of
PT cancer; chronic inflammation and osteoporosis

PS Example 15; Page 44; 64pp: English.

CC The present invention describes peptides (I) that bind to calreticulin
CC (CR), specifically comprising the sequence: KGXIXX3R, where one or more
CC X = basic amino acid (aa). (I) are used to treat cancer (particularly of
CC the prostate or breast, or promyelocytic leukaemia), chronic
CC inflammation (e.g. arthritis) or osteoporosis, also they can be used to
CC treat other bone diseases. (I) act by modulating binding of hormone
CC receptors (HR) to DNA. The present sequence represents a peptide used
CC in the exemplification of the present invention.

XX
SQ Sequence 6 AA:

Query Match 51.5%; Score 17; DB 20; Length 6;
Best Local Similarity 40.0%; Pred. No. 6.4e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 RIFSK 6
Db 1 kvfak 5

Search completed: July 15, 2002, 13:25:25
Job time: 1456 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:26:12 ; Search time 40.89 Seconds
(without alignments)
4.181 Million cell updates/sec

Title: US-09-712-819A-2

Perfect score: 33

Sequence: 1 TRIFSKL 7

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 39160

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/2/1aa/PCITUS.COMB.pep: *
6: /cgn2_6/prodata/2/1aa/Packfiles1.pep: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	23	69.7	7	2	US-09-070-756-9		Sequence 9, Appli
	2	19	57.6	6	1	US-07-634-641-13		Sequence 13, Appli
	3	18	54.5	6	4	US-08-540-922D-6		Sequence 6, Appli
	4	18	54.5	7	4	US-09-461-697-330		Sequence 330, App
	5	17	51.5	5	1	US-07-958-222A-5		Sequence 5, Appli
	6	17	51.5	5	2	US-08-760-075A-5		Sequence 5, Appli
	7	17	51.5	5	4	US-09-338-546-5		Sequence 5, Appli
	8	17	51.5	6	1	US-07-943-709-117		Sequence 117, App
	9	17	51.5	6	1	US-07-943-709-118		Sequence 118, App
	10	17	51.5	7	2	US-08-253-854-7		Sequence 7, Appli
	11	17	51.5	7	2	US-08-392-973A-35		Sequence 35, Appli
	12	17	51.5	7	4	US-09-046-373-10		Sequence 10, Appli
	13	16	48.5	4	1	US-08-419-903A-2		Sequence 2, Appli
	14	16	48.5	4	3	US-09-020-299-4		Sequence 4, Appli
	15	16	48.5	5	1	US-08-118-135A-3		Sequence 3, Appli
	16	16	48.5	5	2	US-08-392-973A-10		Sequence 10, Appli
	17	16	48.5	5	4	US-09-177-249-266		Sequence 266, App
	18	16	48.5	6	2	US-08-392-973A-17		Sequence 17, Appli
	19	16	48.5	6	2	US-08-392-973A-18		Sequence 18, Appli
	20	16	48.5	6	4	US-09-106-216-57		Sequence 57, Appli
	21	16	48.5	7	1	US-08-050-232-9		Sequence 9, Appli
	22	16	48.5	7	1	US-08-457-274A-13		Sequence 13, Appli
	23	16	48.5	7	2	US-08-661-767-9		Sequence 9, Appli
	24	16	48.5	7	2	US-08-392-973A-23		Sequence 23, Appli
	25	16	48.5	7	2	US-08-392-973A-24		Sequence 24, Appli
	26	16	48.5	7	2	US-08-392-973A-25		Sequence 25, Appli
	27	16	48.5	7	4	US-09-031-902-10		Sequence 10, Appli

28	16	48.5	7	5	PCT-US95-05756-13	Sequence 13, Appli
29	15	45.5	4	1	US-08-450-384-21	Sequence 21, Appli
30	15	45.5	4	6	5175113-1	Patent No. 5175113
31	15	45.5	5	1	US-08-200-900A-33	Sequence 33, Appli
32	15	45.5	5	2	US-08-392-973A-9	Sequence 9, Appli
33	15	45.5	5	2	US-08-392-973A-11	Sequence 11, Appli
34	15	45.5	5	2	US-08-896-605A-3	Sequence 3, Appli
35	15	45.5	5	2	US-08-896-501A-3	Sequence 6, Appli
36	15	45.5	5	2	US-08-483-506A-6	Sequence 13, Appli
37	15	45.5	5	5	PCT-US94-00616-33	Sequence 33, Appli
38	15	45.5	6	1	US-07-718-577-13	Sequence 13, Appli
39	15	45.5	6	1	US-08-188-228-3	Sequence 3, Appli
40	15	45.5	6	1	US-08-332-643-3	Sequence 3, Appli
41	15	45.5	6	1	US-08-332-638-3	Sequence 3, Appli
42	15	45.5	6	1	US-08-424-957-6	Sequence 6, Appli
43	15	45.5	6	2	US-08-392-973A-16	Sequence 16, Appli
44	15	45.5	6	2	US-08-392-973A-19	Sequence 19, Appli
45	15	45.5	6	2	US-07-662-764D-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-09-070-756-9
; Sequence 9, Application US/09070756
; Patent No. 5917012
; GENERAL INFORMATION:
; APPLICANT: Nishikata, Makoto
; TITLE OR INVENTION: NOVEL PEPTIDE DERIVATIVES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID S. RESNICK, DIKE, BRONSTEIN, ROBERTS &
; ADDRESSER: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,756
; FILING DATE: 30 April 1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-070-756-9

Query Match 69.7%; Score 23; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIFSK 6
1:111
Db 2 RVFSK 6

RESULT 2

```

US-07-634-641-13
; Sequence 13, Application US/07634641
; Patent No. 5386011
; GENERAL INFORMATION:
; APPLICANT: Wiedeman, Paul E.
; APPLICANT: Kaval, Megumi
; APPLICANT: Lully, Jay R.
; APPLICANT: Or, Yat-Sun
; APPLICANT: Wagner, Rolf
; TITLE OF INVENTION: Hexa- and Heptapeptide Anaphylatoxin
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Abbott Laboratories
; STREET: One Abbott Park Road
; CITY: No. 5386011th Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,641
; FILING DATE: 19901227
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Janssen, Jerry P.
; REGISTRATION NUMBER: 29,175
; REFERENCE/DOCKET NUMBER: 4934.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (708) 937-4558
; TELEFAX: (708) 937-9556
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note="XAA at position 2 is an
; OTHER INFORMATION: L-cyclohexylalanyl residue"
US-07-634-641-13

Query Match 57.5%; Score 19; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 RIFSKL 7
Db 1 RFXFKL 6

RESULT 3
US-08-540-922D-6
; Sequence 6, Application US/08540922D
; Patent No. 6284476
; GENERAL INFORMATION:
; APPLICANT: Boon-Falleur, Thierry; Bichard, Vincent; Van
; APPLICANT: Pel, Aline; De Plaeen, Etienne; Coullie, Pierre;
; APPLICANT: Renaud Jean-Christophe; Wollet, Thomas; and
; APPLICANT: Leche, Bernard.

```

```

; TITLE OF INVENTION: METHOD OF IDENTIFYING INDIVIDUALS SUFFERING
; FROM A CELLULAR ABNORMALITY SOME OF WHOSE
; TITLE OF INVENTION: ABNORMAL CELLS PRESENT COMPLEXES OF HUMAN
; TITLE OF INVENTION: LEUKOCYTE ANTIGEN TYROSYNASE DERIVED
; TITLE OF INVENTION: PEPTIDES, AND METHODS FOR TREATING SAID
; INDIVIDUALS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESS: Felte & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch 1.44 MB storage diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,922D
; FILING DATE: October 11, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,714
; FILING DATE: 28 April 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/994,928
; FILING DATE: 22 December 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary Anne Schofield
; REGISTRATION NUMBER: 36,669
; REFERENCE/DOCKET NUMBER: LUD 5299.5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6
; TYPE: amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; FEATURE:
; OTHER INFORMATION: SEQ of aa corresponding to nt 1714-1731
; OTHER INFORMATION: nt of SEQ ID NO: 1
US-08-540-922D-6

Query Match 54.5%; Score 18; DB 4; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 RIFSKL 7
Db 1 RFXFKL 6

RESULT 4
US-09-461-697-330
; Sequence 330, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Putnam, Kasumi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING

```

TITLE OF INVENTION: CELL DEATH
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 330
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens
US-09-461-697-330

Query Match 54.5%; Score 18; DB 4; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 IFSKL 7
:|:|:
Db 2 LFSRL 6

RESULT 5
US-07-958-222A-5
Sequence 5, Application US/07958222A
Patent No. 5332668
GENERAL INFORMATION:
APPLICANT: YAMASHITA, TAKASHI
APPLICANT: HIGASHI, SUSUMU
APPLICANT: HIGASHI, TOSHIHIKO
APPLICANT: MACHIDA, HARUO
APPLICANT: IMASAKI, SHINJIRO
APPLICANT: BEPPU, TERUHIKO
TITLE OF INVENTION: PROTEASE WITH LOW THERMOSTABILITY AND RELATED
TITLE OF INVENTION: PRODUCT THEREOF AND METHOD FOR PRODUCING THE
TITLE OF INVENTION: SAME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBE, MARTENS, OLSON AND BEAR
STREET: 620 NEWPORT CENTER DRIVE 16TH FLOOR
CITY: NEWPORT BEACH
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/958,222A
FILING DATE: 19921008
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 3-263878
FILING DATE: 11-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: SATOT14.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX: 183513 KNOBE NPBH
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
US-07-958-222A-5

Query Match 51.5%; Score 17; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 IFSK 6
:|:|:
Db 2 LFSK 5

RESULT 6
US-08-760-075A-5
Sequence 5, Application US/08760075A
Patent No. 5942429

GENERAL INFORMATION:
APPLICANT: KIRSCHBAUM, Bernd
APPLICANT: MOELLNER, Stefan
APPLICANT: BARTLETT, Robert
TITLE OF INVENTION: NOVEL DEAH-BOX PROTEINS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,075A

FILING DATE: 04-DEC-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 19545126.0

FILING DATE: 04-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: GRANADOS, Patricia D.

REGISTRATION NUMBER: 33,683

REFERENCE/DOCKET NUMBER: 18748/309

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-760-075A-5

Query Match 51.5%; Score 17; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 RIFSK 6
:|:|:
Db 1 KIFOK 5

RESULT 7

US-09-338-546-5
Sequence 5, Application US/09338546
Patent No. 6251645

GENERAL INFORMATION:

APPLICANT: KIRSCHBAUM, Bernd

APPLICANT: MOELLNER, Stefan

APPLICANT: BARTLETT, Robert

;; TITLE OF INVENTION: NOVEL DEAH-BOX PROTEINS
;; NUMBER OF SEQUENCES: 38
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/338,546
;; FILING DATE:
;; CLASSIFICATION:
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US/08/760,075
;; FILING DATE: 04-DEC-1996
;; APPLICATION NUMBER: DE 19545126.0
;; FILING DATE: 04-DEC-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: GRANADOS, Patricia D.
;; REGISTRATION NUMBER: 33,683
;; REFERENCE/DOCKET NUMBER: 18748/309
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-09-338-546-5

Query Match 51.5%; Score 17; DB 4; Length 5;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RIFSK 6
DB 1 KIFOK 5

RESULT 8
US-07-943-709-117
;; Sequence 117 Application US/07943709
;; Patent No. 5556762
;; GENERAL INFORMATION:
;; APPLICANT: Pinilla, Clemencia
;; APPLICANT: Appel Jr., Jon R.
;; APPLICANT: Blondelle, Silvie
;; APPLICANT: Dooley, Colette T.
;; APPLICANT: Eichter, Jutta
;; APPLICANT: Houghten, Richard A.
;; TITLE OF INVENTION: SCANNING SYNTHETIC PEPTIDE COMBINATORIAL
;; TITLE OF INVENTION: LIBRARIES: OLIGOPEPTIDE MIXTURE SETS HAVING ONE
;; TITLE OF INVENTION: PREDETERMINED RESIDUE AT A SINGLE, PREDETERMINED
;; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
;; NUMBER OF SEQUENCES: 119
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dressler, Goldsmith, Shore, Sutter &
;; ADDRESSER: Milinow, Ltd.
;; STREET: 180 No. 5556762th Stetson, Suite 4700
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60601

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/943,709
;; FILING DATE: 19920911
;; CLASSIFICATION: 530
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 07/797,551
;; FILING DATE: 19-NOV-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Gamson, Edward P.
;; REGISTRATION NUMBER: 29,381
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (312) 616-5400
;; TELEFAX: (312) 616-5460
;; INFORMATION FOR SEQ ID NO: 117:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 6 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /note= "Xaa is Ac-Met."
;; NAME/KEY: Modified-site
;; LOCATION: 6
;; OTHER INFORMATION: /note= "Xaa is Thr-NH2."
;; US-07-943-709-117

Query Match 51.5%; Score 17; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIF 4
DB 2 TKIF 5

RESULT 9
US-07-943-709-118
;; Sequence 118, Application US/07943709
;; Patent No. 5556762
;; GENERAL INFORMATION:
;; APPLICANT: Pinilla, Clemencia
;; APPLICANT: Appel Jr., Jon R.
;; APPLICANT: Blondelle, Silvie
;; APPLICANT: Dooley, Colette T.
;; APPLICANT: Eichter, Jutta
;; APPLICANT: Houghten, Richard A.
;; TITLE OF INVENTION: SCANNING SYNTHETIC PEPTIDE COMBINATORIAL
;; TITLE OF INVENTION: LIBRARIES: OLIGOPEPTIDE MIXTURE SETS HAVING ONE
;; TITLE OF INVENTION: PREDETERMINED RESIDUE AT A SINGLE, PREDETERMINED
;; TITLE OF INVENTION: METHODS OF MAKING AND USING THE SAME
;; NUMBER OF SEQUENCES: 119
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Dressler, Goldsmith, Shore, Sutter &
;; ADDRESSER: Milinow, Ltd.
;; STREET: 180 No. 5556762th Stetson, Suite 4700
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60601
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/943,709
FILING DATE: 19920911
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/797,551
FILING DATE: 19-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P
REGISTRATION NUMBER: 29,381
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5400
TELEFAX: (312) 616-5460
INFORMATION FOR SEQ ID NO: 118:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Xaa is Ac-Thr."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Xaa is Thr-NH2."
US-07-943-709-118

Query Match 51.5%; Score 17; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIF 4
1 : : :
DB 2 TRIF 5

RESULT 10
US-08-253-854-7
Sequence 7, Application US/08253854
Patent No. 5504190
GENERAL INFORMATION:
APPLICANT: Houghten, Richard A.
APPLICANT: Cuervo, Julio H.
APPLICANT: Pinilla, Clemencia
APPLICANT: Appel Jr., Jon R.
APPLICANT: Blondelle, Sylvie
TITLE OF INVENTION: Synthesis of Equimolar Multiple
TITLE OF INVENTION: Oligomer Mixtures, Especially of Oligopeptide Mixtures
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dressler, Goldsmith, Shore, Sutter &
ADDRESSEE: Milnamow, Ltd.
STREET: 180 No. 5504190th Stetson Avenue, Suite 4700
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,854
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gamson, Edward P.
REGISTRATION NUMBER: 29,381
REFERENCE/DOCKET NUMBER: PRL.0003

TELECOMMUNICATION INFORMATION:
TELEPHONE: (312)616-5400
TELEFAX: (312)616-5460
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-253-854-7

Query Match 51.5%; Score 17; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRIFS 5
1 : : :
DB 3 TRIFS 7

RESULT 11
US-08-392-973A-35
Sequence 35, Application US/08392973A
Patent No. 5843453
GENERAL INFORMATION:
APPLICANT: BEATTIE, James
APPLICANT: HOLDER, Andrew T.
TITLE OF INVENTION: GROWTH HORMONE POTENTIATING MOLECULES
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE AND DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: United States of America
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,973A
FILING DATE: 25-APR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286.301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-526-6000
TELEFAX: 617-526-5000
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..7
OTHER INFORMATION:
OTHER INFORMATION: GROWTH HORMONE"
US-08-392-973A-35

Query Match 51.5%; Score 17; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIF 5
Db 1 SRIFF 5

RESULT 12

US-09-046-373-10
; Sequence 10, Application US/09046373
; Patent No. 6235714
; GENERAL INFORMATION:
; APPLICANT: Sudhir Paul
; APPLICANT: Larry J. Smith
; APPLICANT: Gennady Golodov
; TITLE OF INVENTION: Methods for Identifying Inducers and
; TITLE OF INVENTION: Inhibitors of Catalytic Antibodies, Compositions and their
; FILE REFERENCE: UNMC 63123
; CURRENT APPLICATION NUMBER: US/09/046,373
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 7
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-046-373-10

Query Match 51.5%; Score 17; DB 4; Length 7;
Best Local Similarity 40.0%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 IESKL 7
; 1:1:
Db 2 VENKI 6

RESULT 13

US-08-419-903A-2
; Sequence 2, Application US/08419903A
; Patent No. 5753226
; GENERAL INFORMATION:
; APPLICANT: Greene, Mark I.
; APPLICANT: Colasrelli, George
; TITLE OF INVENTION: METHODS OF ENHANCING EPITHELIAL CELL
; TITLE OF INVENTION: PROLIFERATION
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Maciewicz & NO. 5753226R1S
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: United States of America
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,903A
; FILING DATE: 11-Apr-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-1705
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids

TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-419-903A-2

Query Match 48.5%; Score 16; DB 1; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 ESKL 7
; 1:1:
Db 1 FNKL 4

RESULT 14

US-09-020-299-4
; Sequence 4, Application US/09020299
; Patent No. 6090381
; GENERAL INFORMATION:
; APPLICANT: LEUNG, Shui-on
; APPLICANT: OU, Zhengxing
; TITLE OF INVENTION: STIMULATION OF AN IMMUNE RESPONSE WITH ANTIBODIES
; TITLE OF INVENTION: LABELED WITH THE a-GALACTOSYL EPI TOPE
; FILE REFERENCE: 018733/0814
; CURRENT APPLICATION NUMBER: US/09/020,299
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: US 60/037,908
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 4
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-020-299-4

Query Match 48.5%; Score 16; DB 3; Length 4;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIF 4
; 1:1:
Db 1 SRIFF 4

RESULT 15

US-08-118-135A-3
; Sequence 3, Application US/08118135A
; Patent No. 5437982
; GENERAL INFORMATION:
; APPLICANT: Catterall, W.A. and Ebnoltz, G.
; TITLE OF INVENTION: Specific Inactivation Gate Inhibitors of the Sodium Channe
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Christensen, O'Connor, Johnson and Kindness
; STREET: 2800 Pacific First Center, 1420 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101-2347
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette-5.25 inch, 1.2mb storage
; COMPUTER: IBM PC/386 Compatible
; OPERATING SYSTEM: MS-DOS 4.01
; SOFTWARE: Word 5.5-L
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,135A
; FILING DATE: September 8, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: NO. 5437982e
; ATTORNEY/AGENT INFORMATION:

NAME: Broderick, Thomas F.
 REGISTRATION NUMBER: 31,332
 REFERENCE/DOCKET NUMBER: UOFW16993
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 1-206-682-8100; 1-206-224-0709 (direct)
 TELEFAX: 1-206-224-0779
 TELEX: 4938023
 INFORMATION FOR SEQ. ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 DESCRIPTION: page 5, line 4; KIFMK
 US-08-118-135A-3

Query Match 48.5%; Score 16; DB 1; Length 5;
 Best Local Similarity 60.0%; Pred. No. 1.7e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RIFSK 6
 : 1 1
 Db 1 KIFMK 5

Search completed: July 15, 2002, 13:26:13
 Job time: 493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:27:08 ; Search time 49.13 Seconds

(without alignments)
13.691 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35

Sequence: 1 FYOLALV 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR-71.*

1: pirl.*

2: pirl2.*

3: pirl3.*

4: pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	13	37.1	3	568328	blood cell protein
2	13	37.1	5	B44823	synaposomal-assoc
3	13	37.1	7	E30608	Ig kappa chain V-I
4	13	37.1	7	E33932	Ig mu chain D regi
5	12	34.3	5	140469	dnazX-like protein
6	12	34.3	7	B39127	phosphotransferase
7	11	31.4	6	A37765	hypothetical prote
8	11	31.4	7	S19630	ribosomal protein
9	10	28.6	4	S18401	thyroglobulin - do
10	10	28.6	4	A32039	tyrosine-melanocyt
11	10	28.6	4	J01273	neuropeptide Antho
12	10	28.6	6	S60293	tubulin beta-3 cha
13	10	28.6	6	B35640	cerebellar degener
14	10	28.6	6	149421	lamtalin B1 - weste
15	10	28.6	6	A43129	neuropeptide GNPFR
16	10	28.6	7	S55548	mcrb protein - Esc
17	10	28.6	7	PN0150	omega-gliadin 1'
18	10	28.6	7	PS0254	18K protein 5507 -
19	10	28.6	7	E48394	glycoprotein compo
20	10	28.6	7	B48394	major fat-globule
21	9	25.7	4	138888	COI intron 16 prot
22	9	25.7	5	A60521	glycogen phosphory
23	9	25.7	5	S68328	blood cell protein
24	9	25.7	6	148126	alpha-tubulin - Ch
25	9	25.7	7	S33246	neuromodulatory pe
26	9	25.7	7	A38081	amine oxidase (cop
27	8	22.9	3	A22565	R-phycoerythrin al
28	8	22.9	4	A37832	phenol 2-monooxyge
29	8	22.9	4	PT0697	T-cell receptor be

30	8	22.9	5	2	JN0860	peptidyl-dipeptida
31	8	22.9	5	2	E60274	major protein anti
32	8	22.9	5	2	P22565	R-phycoerythrin ga
33	8	22.9	5	2	P80324	ribulose-bisphospha
34	8	22.9	5	2	S55237	zinc-binding prote
35	8	22.9	5	2	JH0253	gut pentapeptide -
36	8	22.9	5	2	PT0590	T-cell receptor be
37	8	22.9	5	2	B44817	34.5k structural p
38	8	22.9	6	2	D44817	35k structural pro
39	8	22.9	6	2	S11556	hydrogensulfite re
40	8	22.9	6	2	A43766	28k ubiquitin-immu
41	8	22.9	6	2	I37027	protamine PI - gor
42	8	22.9	6	2	B33932	Ig mu chain D regi
43	8	22.9	6	2	A41946	T-cell receptor ga
44	8	22.9	7	1	NYPG7	hypothalamic hepta
45	8	22.9	7	2	S21230	dermorphin (Trp-4,

ALIGNMENTS

RESULT 1
S68328
blood cell protein A - Molgula manhattensis (fragment)
C:Species: Molgula manhattensis
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: S68328
R:Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A:Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from
A:Reference number: S68325; MUID:96132650
A:Accession: S68328
A:Molecule type: protein
A:Residues: 1-3 <TAY>

Query Match 37.1%; Score 13; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FY 2
Db 2 FY 3

RESULT 2
B44823
synaposomal-associated protein SNAP-25 peptide 10A - rabbit (fragment)
N:Alternate names: superprotein peptide 10A
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 15-Jun-1996
C:Accession: B44823
J:Loewy, A.; Liu, W.S.; Baillinger, C.; Willard, M.B.
R. Neurosci. 11, 3412-3421, 1991
A:Title: The major 35S-methionine-labeled rapidly transported protein (superprotein)
A:Reference number: A44823; MUID:92044785
A:Accession: B44823
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-5 <LOE>
A:Experimental source: visual tissue
A:Note: sequence extracted from NCBI backbone (NCBI:64255)
C:Keywords: membrane trafficking

Query Match 37.1%; Score 13; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QVAL 6
Db 2 QMAI 5

RESULT

3

E30608 Ig kappa chain V-III region (Cag) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996

C:Accession: E30608

R:Contl. F.R.; Chen, P.P.; McGinnis, D.; Arjona, M.L.; Fernandez, J.; Carson, D.; Sold

J. Immunol. 142, 3158-3163, 1989

A:Title: Structural and idiotype characterization of the L chains of human IgM autoanti

A:Reference number: A30601; MUID:89215279

A:Accession: E30608

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-7 <GON>

C:Keywords: heterotrimer; immunoglobulin

Query Match

Best Local Similarity 37.1%; Score 13; DB 2; Length 7;

Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QVALT 7

DB 1 EIVLT 5

RESULT

4

E33932 Ig mu chain D region (E7) - mouse

C:Species: Mus musculus (house mouse)

C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996

C:Accession: E33932

R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.

Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989

A:Title: Two murine natural polyreactive autoantibodies are encoded by nonmutated germ-l

A:Reference number: A33932; MUID:89282823

A:Accession: E33932

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-7 <BAC>

A:Cross-references: GB:M27106

C:Keywords: Immunoglobulin

Query Match

Best Local Similarity 37.1%; Score 13; DB 2; Length 7;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FY 2

DB 3 FY 4

RESULT

5

I40469 dnanx-like protein - Bacillus subtilis (fragment)

C:Species: Bacillus subtilis

C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000

C:Accession: I40469

R:Struck, J.C.; Hartmann, R.K.; Toschka, H.Y.; Erdmann, V.A.

Mol. Gen. Genet. 215, 478-482, 1989

A:Title: Transcription and processing of Bacillus subtilis small cytoplasmic RNA.

A:Reference number: I40469; MUID:89218958

A:Accession: I40469

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5 <RES>

A:Cross-references: EMBL:X14796; NID:940130; PIDN:CAA32902.1; PID:94376204

C:Genetics:

A:Start codon: GTG

Query Match

Best Local Similarity 34.3%; Score 12; DB 2; Length 5;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YO 3

DB 3 YO 4

RESULT

6

B39127 phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)

C:Species: Escherichia coli

C>Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999

C:Accession: B39127

R:Hardisty, C.; Ferran, C.; Di Rienzo, J.M.

J. Bacteriol. 173, 449-456, 1991

A:Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of

A:Reference number: A39127; MUID:91100329

A:Accession: B39127

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-7 <HAR>

A:Cross-references: GB:M38416; NID:9155142; PIDN:AAA98418.1; PID:9155144

C:Keywords: phosphotransferase

Query Match

Best Local Similarity 34.3%; Score 12; DB 2; Length 7;

Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYOLA 5

DB 3 FEQIS 7

RESULT

7

A37765 hypothetical protein (csma 5' region) - Chloroflexus aurantiacus (fragment)

C:Species: Chloroflexus aurantiacus

C>Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993

C:Accession: A37765

R:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.

J. Bacteriol. 172, 4497-4504, 1990

A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus auranti

A:Reference number: A37765; MUID:90330558

A:Accession: A37765

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-6 <THE>

A:Cross-references: GB:M33964

Query Match

Best Local Similarity 31.4%; Score 11; DB 2; Length 6;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 QLA 5

DB 2 QMA 4

RESULT

8

S19630 ribosomal protein L30 - Streptomyces griseus (fragment)

C:Species: Streptomyces griseus

C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1997

C:Accession: S19630

R:Oehl, K.

Int. J. Syst. Bacteriol. 42, 144-150, 1992

A:Title: Electrophoretic heterogeneity of ribosomal protein AT-L30 among actinomycete

A:Reference number: S19630; MUID:92144363

A:Accession: S19630
 A:Molecule type: protein
 A:Residues: 1-7 <CCH>
 A:Experimental source: strain IFO 13189
 C:Superfamily: Escherichia coli ribosomal protein L30
 C:Keywords: protein biosynthesis; ribosome

Query Match 31.4%; Score 11; DB 2; Length 7;
 Best Local Similarity 40.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 QLALT 7
 : 1
 Db 2 RLKLT 6

RESULT 9

S18401
 thyroglobulin - dog (fragment)
 C:Species: Canis lupus familiaris (dog)
 C:Date: 19-Mar-1997 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
 C:Accession: S18401
 R:Donat, A.; Vassart, G.; Christophe, D.
 Biochim. Biophys. Acta 1090, 235-237, 1991
 A:Title: Isolation and characterization of the canine thyroglobulin gene promoter region
 A:Reference number: S18401; MUID:92031697
 A:Accession: S18401
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-4 <DON>
 A:Cross-references: GB:S61184; NID:g237714; PID:NAB20127.1; PID:g237715
 C:Superfamily: thyroglobulin; cholinesterase homology; thyroglobulin type I repeat homol
 C:Keywords: duplication; iodine; thyroid gland; thyroid hormone biosynthesis

Query Match 28.6%; Score 10; DB 2; Length 4;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LAL 6
 : 1
 Db 1 MAL 3

RESULT 10

A32039
 tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000
 C:Accession: A32039
 R:Horvath, A.; Kastin, A.J.
 J. Biol. Chem. 264, 2175-2179, 1989
 A:Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting factor
 A:Reference number: A32039; MUID:99133285
 A:Accession: A32039
 A:Molecule type: protein
 A:Residues: 1-4 <HOR>
 A:Experimental source: brain
 C:Superfamily: unassigned animal peptides
 C:Keywords: amidated carboxyl end
 F:4/Modified site: amidated carboxyl end (gly) #status experimental

Query Match 28.6%; Score 10; DB 2; Length 4;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YOL 4
 : 1
 Db 1 YPL 3

RESULT 11

J01273
 neuropeptide Antho-Kamide - sea anemone (Anthopleura elegantissima)
 C:Species: Anthopleura elegantissima
 C:Date: 31-Mar-1992 #sequence_revision 04-Dec-1992 #text_change 08-Dec-1995
 C:Accession: J01273
 R:Notack, H.P.; Rinehart, K.L.; Grimmelikhuijzen, C.J.P.
 Biochem. Biophys. Res. Commun. 179, 1205-1211, 1991
 A:Title: Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a novel neuro
 A:Reference number: J01273; MUID:92028852
 A:Accession: J01273
 A:Molecule type: protein
 A:Residues: 1-4 <NOT>
 C:Comment: The carboxyl-terminal amide probably arises from cleavage of a following g
 C:Keywords: amidated carboxyl end; neuropeptide; phenyllactylation
 F:1/Modified site: L-3-phenyllactic acid (Phe) #status experimental
 F:4/Modified site: amidated carboxyl end (Ala) #status experimental

Query Match 28.6%; Score 10; DB 2; Length 4;
 Best Local Similarity 33.3%; Pred. No. 2.8e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYO 3
 : 1
 Db 1 FFK 3

RESULT 12

S60293
 tubulin beta-3 chain - fruit fly (Drosophila melanogaster) (fragment)
 C:Species: Drosophila melanogaster
 C:Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 17-Mar-1999
 C:Accession: S60293
 R:Chapel, S.; Sobrier, M.L.; Montpied, P.; Micard, D.; Brubhat, A.; Couderc, J.L.; Das
 Insect Mol. Biol. 2, 39-48, 1993
 A:Title: In Drosophila Kc cells 20-OHE induction of the 60C beta-3 tubulin gene expe
 A:Reference number: S60292; MUID:9724543
 A:Accession: S60293
 A:Molecule type: DNA
 A:Residues: 1-6 <CHA>
 A:Cross-references: EMBL:X60394
 C:Genetics:
 A:Gene: FlyBase:fbgtr
 A:Cross-references: FlyBase:FBgn0003888
 A:Introns: 4/1

Query Match 28.6%; Score 10; DB 2; Length 6;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ALT 7
 : 1
 Db 4 AVT 6

RESULT 13

B35640
 cerebellar degeneration-related protein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993
 C:Accession: B35640
 R:Chen, Y.T.; Rettig, W.J.; Yenamandra, A.K.; Kozak, C.A.; Chaganti, R.S.K.; Posner,
 Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990
 A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal
 A:Reference number: A35640; MUID:90222173
 A:Accession: B35640
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-6 <CHE>

Query Match 28.6%; Score 10; DB 2; Length 6;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0;

QY 1 FYQ 3
1:
DB 1 FWE 3

RESULT 14

I49421
Iaminin B1 - western wild mouse (fragment)
C:Species: Mus spretus (western wild mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49421
R:KO, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maizaki, Y.; Nadeau, J.H.
Mamm. Genome 5, 349-355, 1994
A>Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48334; MUID:94319082
A:Accession: I49421
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6 <RES>
A:Cross-references: EMBL:U05736; NID:q497073; PIDN:AA60477.1; PID:9642829

Query Match 28.6%; Score 10; DB 2; Length 6;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0;

QY 2 YQLAL 6
1:
DB 2 YSTCL 6

RESULT 15

AA3129
neuropeptide GnFRFamide - tapeworm (Moniezia expansa)
C:Species: Moniezia expansa
C>Date: 10-Nov-1997 #sequence_revision 14-Nov-1997 #text_change 14-Nov-1997
C:Accession: AA3129
R:Maule, A.; Shaw, C.; Halton, D.; Thim, L.
Biochem. Biophys. Res. Commun. 193, 1054-1060, 1993
A>Title: GnFRFamide: A novel FMRFamide-immunoreactive peptide isolated from the sheep t
A:Reference number: AA3129; MUID:93312289
A:Accession: AA3129
A:Molecule type: protein
A:Residues: 1-6 <MAU>
C:Keywords: amidated carboxyl end; neuropeptide
P:6/Modified site: amidated carboxyl end (Phe) #status predicted

Query Match 28.6%; Score 10; DB 2; Length 6;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0;

QY 1 FYQ 3
1:
DB 3 FFR 5

Search completed: July 15, 2002, 13:27:08
Job time: 438 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:38:10 ; Search time 24.49 Seconds

(without alignments)
11.067 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35
Sequence: 1 FYOLALF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	45.7	4	1	FYRI_ANTLR
2	12	34.3	5	1	UC22_MAIZE
3	11	31.4	5	1	AL14_CARMA
4	11	31.4	7	1	AL17_CYPPO
5	11	31.4	7	1	GERP_MOUSE
6	10	28.0	4	1	FFRK_ANTLR
7	10	28.6	6	1	FARP_MONEX
8	10	28.6	7	1	AL12_CARMA
9	10	28.6	7	1	AL13_CARMA
10	10	28.6	7	1	AL14_CARMA
11	10	28.6	7	1	AL15_CARMA
12	9	25.7	5	1	PAP2_PARMA
13	9	25.7	5	1	RE11_LITRU
14	9	25.7	5	1	RE21_LITRU
15	9	25.7	5	1	RE31_LITRU
16	9	25.7	5	1	RE32_LITRU
17	9	25.7	7	1	E105_LITRU
18	9	25.7	7	1	TY51_LITRU
19	9	25.7	7	1	TY51_LITRU
20	9	25.7	7	1	TY51_LITRU
21	8	22.9	6	1	MM22_ACHFU
22	8	22.9	7	1	C1A_ENTFA
23	8	22.9	7	1	HY7_PIG
24	8	22.9	7	1	MM21_LEPDE
25	8	22.9	7	1	FARP_HIRME
26	7	20.0	4	1	FARP_HIRME
27	7	20.0	4	1	RM01_YEAST
28	7	20.0	5	1	B104_CITTR
29	7	20.0	5	1	E104_LITRU
30	7	20.0	5	1	FARP_ANTLR
31	7	20.0	5	1	PRCT_PBRAM
32	7	20.0	6	1	PSK_DAVCA
33	7	20.0	6	1	AS22_LACSN
			7	1	CARP_MYTEO

34	7	20.0	7	1	FAR2_ASCSU	P31890 ascaris suu
35	7	20.0	7	1	FAR3_HAECCO	P81298 haemochus
36	7	20.0	7	1	FAR3_PANRE	P41874 panagrellus
37	7	20.0	7	1	FAR5_HIRME	P42564 hirudo medi
38	7	20.0	7	1	UH11_RAF	P56575 rattus norv
39	7	20.0	7	1	UN05_PINPS	P81675 rattus norv
40	6	17.1	4	1	ACH1_ACHFU	P35904 achina fu
41	6	17.1	4	1	DCMS_PSECH	P19918 pseudomonas
42	6	17.1	4	1	FLRF_HIRME	P42561 hirudo medi
43	6	17.1	4	1	FLRN_ANTLR	P58707 anthopleura
44	6	17.1	4	1	FMRE_MACNT	P01162 macrocallis
45	6	17.1	4	1	OCPI_OCCMI	P58648 octopus min

ALIGNMENTS

RESULT	1	ALIGNMENTS
FYRI_ANTLR		
ID	FYRI_ANTLR	STANDARD: PRT: 4 AA.
AC	P58706:	
DT	01-MAR-2002 (Rel. 41, Created)	
DT	01-MAR-2002 (Rel. 41, Last sequence update)	
DT	01-MAR-2002 (Rel. 41, Last annotation update)	
DE	Antho-Riamide I [Contains: Antho-Riamide II].	
OS	Anthopleura elegantissima (Sea anemone).	
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinoptaria;	
OC	Nyanthaea; Actinoptaria; Anthopleura.	
OX	NCBI_TaxID=6110;	
RN	[1]	
RP	SEQUENCE.	
RP	Pubmed=1821096;	
RA	Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,	
RA	Grimmelikhuizen C.J.P.;	
RT	*Isolation of two novel neuropeptides from sea anemones: the unusual,	
RT	des-phenylalanyl fragment Tyr-Arg-Ile-NH ₂ and its	
RT	Peptides 12:1165-1173(1991).	
RN	[2]	
RP	FUNCTION.	
RX	Pubmed=8397415;	
RA	McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuizen C.J.P.;	
RT	*The expansion behaviour of sea anemones may be coordinated by two	
RT	inhibitory neuropeptides, Antho-Kamide and Antho-Riamide-2.	
RL	Proc. R. Soc. Lond., B. Biol. Sci. 253:183-188(1993).	
CC	-I- FUNCTION: Inhibits spontaneous contractions in several muscle	
CC	groups. May be involved in the expansion phase of feeding	
CC	behaviour in sea anemones.	
CC	-I- SUBCELLULAR LOCATION: Secreted.	
CC	-I- TISSUE SPECIFICITY: Neuron-specific.	
KW	Neuropeptide; Amidation.	
FT	CHAIN 1	ANTHO-RIAMIDE I.
FT	CHAIN 2	ANTHO-RIAMIDE II.
FT	MOD_RES 1	L-3-PHENYLACTYL.
FT	MOD_RES 4	AMIDATION.
SQ	SEQUENCE 4 AA: 598 MW: 60441B59A0000000 CRC64:	
Query Match 45.7% Score 16: DB 1: Length 4:		
Best Local Similarity 50.0%: Pred. No. 1e+05: Indels 0: Gaps 0:		
Matches 2: Conservative 2: Mismatches 0: Indels 0: Gaps 0:		
QY	1 FYOL 4	
DB	1 FYRI 4	
RESULT 2		
UC22_MAIZE	STANDARD: PRT: 5 AA.	
ID	UC22_MAIZE	
AC	P80628:	
DT	01-OCT-1996 (Rel. 34, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	

DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 474)
 DE (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RC SEQUENCE.
 RP Tissue=Coleoptile;
 RA Toulzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Perrotlet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program."
 RL Theor. Appl. Genet. 93:197-1005(1996).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.1, ITS MW IS: 30.4 KDa.
 DR MaizeDB: 123954; COLEOPTILE.
 FT NON-TER 1 1
 FT MOD_RES 5 5
 SO SEQUENCE 5 AA; 654 MW; 72CB19C9C0300000 CRC64;

Query Match 34.3%; Score 12; DB 1; Length 5;
 Best Local Similarity 25.0%; Pred. No. 1e+05;
 Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YQAL 4
 DB 2 YFGL 5

RESULT 3
 AL14_CARMA STANDARD; PRT: 5 AA.
 ID AL14_CARMA
 AC P81817;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinus maenas 14.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC Tissue=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas."
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation; Multigene family.
 DR MOD_RES 5 5
 FT INIT_MET 5
 FT SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 5;
 Best Local Similarity 40.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YQAL 6
 DB 1 YFGL 5

RESULT 4

AL17_CYPDO STANDARD; PRT: 7 AA.
 ID AL17_CYPDO
 AC P82158;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydista 7.
 OS Cydista pomonella (Codling moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
 OX NCBI_TaxID=82600;
 RN [1]
 RC SEQUENCE.
 RP Tissue=Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily."
 RL Peptides 18:1301-1309(1997).
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation.
 DR MOD_RES 7 7
 FT INIT_MET 7
 FT SEQUENCE 7 AA; 873 MW; 672879CABB569350 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 7;
 Best Local Similarity 40.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YQAL 6
 DB 3 YDFGL 7

RESULT 5
 GFRP_MOUSE STANDARD; PRT: 7 AA.
 ID GFRP_MOUSE
 AC P99025;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP cyclonucleoside I feedback regulatory protein (P35) (Fragment).
 GN GCHFR OR GFRP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC Tissue=Liver;
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Van J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Cowthorne N.G. (Aug-1998) to the SWISS-PROT data bank.
 RL Submitted (Aug-1998) to the SWISS-PROT data bank.
 CC -1- FUNCTION: MEDIATES TETRAHYDROBIOPTERIN INHIBITION OF GTP
 CC CYCLOHYDROLYASE I. THIS INHIBITION IS REVERSED BY L-PHENYLANILINE
 CC (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC SWISS-2DPAGE: P99025; MOUSE.
 DR INIT_MET 0
 FT NON-TER 7 7
 FT SEQUENCE 7 AA; 806 MW; 71B5B057273B4700 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 7;
 Best Local Similarity 50.0%; Pred. No. 1e+05;
 Matches 3; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YQAL 7
 DB 2 YLLIST 7

```

RESULT 6
FFKA_AMEL STANDARD: PRT: 4 AA.
AC P58705;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Antho-Kamide.
OC Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nynanthaeae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN
RP SEQUENCE.
RX PubMed=1681803;
RA Nothacker H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenylacetyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a
RT novel neuropeptide from sea anemones."
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN
RP FUNCTION.
RX PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Kamide and Antho-Ramide."
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -1- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Neuron-specific.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
FT SEQUENCE 4 AA: 512 MW: 6D339C9A000000 CRC64;
SQ

Query Match 28.6%; Score 10; DB 1; Length 4;
Best Local Similarity 33.3%; Pred. No. 1e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQ 3
DB 1 FFK 3

RESULT 7
FARP_MONEX STANDARD: PRT: 6 AA.
ID FARP_MONEX
AC P41966;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FMRFamide-like neuropeptide GNFRR-amide.
OS Moniezia expansa (Sheep tapeworm).
OC Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
OC Rhabdliophora; Eutelethophora; Reverteospermatia; Mediofusata;
OC Neodermata; Cestoda; Eucestoda; Cyclophyllidae; Anoplocephalidae;
OC Moniezia.
OX NCBI_TaxID=28841;
RN
RP SEQUENCE.
RX MEDLINE=93312289; PubMed=8323531;
RA Maule A.G., Shaw C., Hulton D.W., Thim L.;
RT "GFRRFamide: a novel FMRFamide-immunoreactive peptide isolated from
RT the sheep tapeworm, Moniezia expansa."
RL Biochem. Biophys. Res. Commun. 193:1054-1060(1993).
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 6 6 AMIDATION.
FT SEQUENCE 6 AA: 787 MW: 69D409C9C4481000 CRC64;
SQ

```

```

Query Match 28.6%; Score 10; DB 1; Length 6;
Best Local Similarity 33.3%; Pred. No. 1e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQ 3
DB 3 FFR 5

RESULT 8
ALL2_CARMA STANDARD: PRT: 7 AA.
ID ALL2_CARMA
AC P81805;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 2.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eubranchyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN
RP SEQUENCE.
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7 7 AMIDATION (POTENTIAL).
FT SEQUENCE 7 AA: 770 MW: 672879CDB5DB70 CRC64;
SQ

Query Match 28.6%; Score 10; DB 1; Length 7;
Best Local Similarity 40.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 YQAL 6
DB 3 YAFGL 7

RESULT 9
ALL3_CARMA STANDARD: PRT: 7 AA.
ID ALL3_CARMA
AC P81806;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 3.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubranchyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN
RP SEQUENCE.
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.

```

KM Neuropeptide: Multigene family.
SQ SEQUENCE 7 AA; 796 MW; 672879CDCB476B70 CRC64;

Query Match 28.6%; Score 10; DB 1; Length 7;
Best Local Similarity 40.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YOLAL 6
1
3 YAFGL 7

RESULT 10
ALL4_CARMA STANDARD; PRT; 7 AA.
ID ALL4_CARMA
AC P81807;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas (Common shore crab) (Green crab).
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;

*Isolation and identification of multiple neuropeptides of the RT allatostatin superfamily in the shore crab Carcinus maenas.;"
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUOTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Multigene family.
SQ SEQUENCE 7 AA; 782 MW; 672879CDCB476AC0 CRC64;

Query Match 28.6%; Score 10; DB 1; Length 7;
Best Local Similarity 40.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YOLAL 6
1
3 YAFGL 7

RESULT 11
ALL5_CARMA STANDARD; PRT; 7 AA.
ID ALL5_CARMA
AC P81808;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas (Common shore crab) (Green crab).
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;

*Isolation and identification of multiple neuropeptides of the RT allatostatin superfamily in the shore crab Carcinus maenas.;"
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUOTRANSMITTER OR NEUROMODULATOR.

CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide: Multigene family.
FT MOD_RES 7 7 AMIDATION
SQ SEQUENCE 7 AA; 781 MW; 672879CDCB476420 CRC64;

Query Match 28.6%; Score 10; DB 1; Length 7;
Best Local Similarity 40.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 YOLAL 6
1
3 YAFGL 7

RESULT 12
PAP2_PARMA STANDARD; PRT; 5 AA.
ID PAP2_PARMA
AC P81864;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pardaxin II (PXII) (Fragment).
OS Pardachirus marmoratus (Red sea moose sole).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Pleuronectiformes;
OC Soleiidae; Soleidae; Pardachirus.
OX NCBI_TaxID=31087;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RX MEDLINE=87057369; PubMed=3782138;
RA Lazarovici P., Primor N., Loew L.M.;

*Purification and pore-forming activity of two hydrophobic RT polypeptides from the secretion of the Red sea moose sole (Pardachirus RT marmoratus).;"
RL J. Biol. Chem. 261:16704-16713(1986).
CC -1- FUNCTION: EXHIBITS UNUSUAL SHARK REPELLENT AND SURFACTANT PROPERTIES. FORMS VOLTAGE-DEPENDENT, ION-PERMEABLE CHANNELS IN MEMBRANES. AT HIGH CONCENTRATION CAUSES CELL MEMBRANE LYSIS.
CC -1- SUBUNIT: MONOMER. IN AQUEOUS SOLUTION EXISTS AS A Tetramer.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE PARDAXIN FAMILY.
KM Toxin.
FT NON_TER 5 5
SQ SEQUENCE 5 AA; 614 MW; 7769C9C9C8100000 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
1
2 FF 3

RESULT 13
RELI_LITRU STANDARD; PRT; 5 AA.
ID RELI_LITRU
AC P82070;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellididn 1.1.
OS Utiroia rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidae; Hylidae;
OC Utiroia.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.

RC TISSUE-Skin secretion;
 RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
 RA Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 'Litoria rubella', the skin peptide profile as a probe for the study
 of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=598; METHOD=FAB.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CBZ00000 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 5;
 Best Local Similarity 50.0%; Pred. No. 1e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
 I;
 Db 3 FF 4

RESULT 14
 RE21_LITRU STANDARD; PRT; 5 AA.
 ID RE21_LITRU
 AC P82071;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 2.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_TaxID-104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE-Skin secretion;
 RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
 RA Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 'Litoria rubella', the skin peptide profile as a probe for the study
 of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=626; METHOD=FAB.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 5;
 Best Local Similarity 50.0%; Pred. No. 1e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
 I;
 Db 3 FF 4

RESULT 15
 RE31_LITRU STANDARD; PRT; 5 AA.
 ID RE31_LITRU
 AC P82072;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 3.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_TaxID-104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE-Skin secretion;
 RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
 RA Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 'Litoria rubella', the skin peptide profile as a probe for the study
 of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAB.
 KW Amphibian skin; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 5;
 Best Local Similarity 50.0%; Pred. No. 1e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
 I;
 Db 3 FF 4

Search completed: July 15, 2002, 13:38:10
 Job time: 710 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:37:39 ; Search time 86.73 Seconds
(without alignments)
13.962 Million cell updates/sec

Title: US-09-712-819A-3
Perfect score: 35
Sequence: 1 FYQLAIR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 65

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	34.3	7	8 099182	099182 gnatholebia
2	10	28.6	7	7 P72081	P72081 nocardia la
3	9	25.7	5	13 P82070	P82070 litorea rub
4	9	25.7	5	13 P82071	P82071 litorea rub
5	9	25.7	5	13 P82072	P82072 litorea rub
6	9	25.7	5	13 P82073	P82073 litorea rub
7	9	25.7	7	12 09Y010	09Y010 transmissid
8	9	25.7	7	13 P82065	P82065 litorea rub
9	9	25.7	7	13 P82101	P82101 litorea rub
10	8	22.9	7	8 095945	095945 saccharomyc
11	8	22.9	7	10 09C5B3	09C5B3 arabidopsis
12	8	22.9	7	11 063668	063668 rattus norv
13	8	22.9	7	12 066113	066113 cherry leaf
14	7	20.0	5	13 P82100	P82100 litorea rub
15	7	20.0	6	10 P82181	P82181 spinacia ol
16	7	20.0	6	10 P82182	P82182 spinacia ol

17	7	20.0	7	2 007354	007354 synechococc
18	7	20.0	7	2 047029	047029 enterobacte
19	7	20.0	7	2 050556	050556 actinobacti
20	7	20.0	7	2 034028	034028 sphingomon
21	7	20.0	7	2 054248	054248 streptomyce
22	7	20.0	7	4 015897	015897 homo sapien
23	7	20.0	7	10 P93233	P93233 lycopersico
24	7	20.0	7	12 09Y1R0	09Y1R0 human adeno
25	7	20.0	7	12 09Y1R0	09Y1R0 human adeno
26	7	20.0	7	12 09Y1R3	09Y1R3 human adeno
27	6	17.1	5	10 09Y007	09Y007 hordeum vul
28	6	17.1	5	13 P82099	P82099 litorea rub
29	6	17.1	6	13 P82096	P82096 litorea rub
30	6	17.1	7	4 015903	015903 homo sapien
31	6	17.1	7	8 098866	098866 spinacia ol
32	6	17.1	7	10 P82445	P82445 nicotiana t
33	6	17.1	7	13 042564	042564 fugu rubrip
34	6	17.1	7	15 007624	007624 fugu sarcom
35	5	14.3	5	2 P83073	P83073 bacillus ce
36	5	14.3	6	10 P82541	P82541 spinacia ol
37	5	14.3	7	2 047477	047477 escherichia
38	5	14.3	7	2 047505	047505 escherichia
39	5	14.3	7	2 P70804	P70804 azotobacter
40	5	14.3	7	6 028742	028742 oryctolagus
41	5	14.3	7	8 P92421	P92421 psathyrosta
42	5	14.3	7	8 P92385	P92385 hordeum mar
43	5	14.3	7	8 P92372	P92372 haynaldia v
44	5	14.3	7	8 P92403	P92403 lophopyrum
45	5	14.3	8	8 P92425	P92425 pseudoroegn

ALIGNMENTS

RESULT 1
AC 099182 PRELIMINARY; PRT; 7 AA.
ID 099182:
DT 01-MAY-1999 (TREMBLER, 10, Created)
DT 01-MAY-1999 (TREMBLER, 10, Last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
DE CYTOCHROME OXIDASE I (FRAGMENT).
GN COI.
OS Gnatholebias zonatus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Gnatholebias.
OX NCBI_TaxID=135316;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20072928; PubMed=10603257;
RA Murphy W.J., Thomerson J.E., Collier G.E.;
RT "Phylogeny of the Neotropical killifish family Rivulidae
(Cyprinodontiformes, Aplocheilidae) inferred from mitochondrial DNA
sequences.";
RL Mol. Phylogenet. Evol. 13:269-301(1999).
DR EMBL; AF002591; AAD01074.1; -.
KW Mitochondrion.
FT NON_TER
SQ SEQUENCE 7 AA; 899 MW; 6272121P6CB572030 CRC64;

Query Match 34.3%; Score 12; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 Y0 3
11
Db 3 Y0 4

```

RESULT 2
ID P72081 PRELIMINARY; PRT; 7 AA.
AC P72081.
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 3'-METHYLCEPHEM HYDROXYLASE (FRAGMENT).
GN CEFEP.
OS Nocardiaceae.
OC Actinomycetales; Pseudonocardiales; Pseudonocardaceae; Amycolatopsis.
OC NCB1_TaxID=1913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96009872; PubMed=7557411;
RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
RA Liras P.;
RT "Characterization of the cmcH genes of Nocardiia lactamurans and
RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
RT O-carbamoyltransferase for cephamycin biosynthesis.";
RL Gene 162:21-27(1995).
DR EMBL; Z21682; CAA79797.1; -.
FT NON_TER 1
FT SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

```

```

Query Match 28.6%; Score 10; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 5 ALT 7
DB 4 AVT 6

RESULT 3
ID P82070 PRELIMINARY; PRT; 5 AA.
AC P82070.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELLIDIN 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCB1_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=598; METHOD=FAB.
KM Amphibian skin.
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

```

```

Query Match 25.7%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 FY 2
DB 3 FF 4

```

```

RESULT 4
ID P82071 PRELIMINARY; PRT; 5 AA.
AC P82071.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELLIDIN 2.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCB1_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=626; METHOD=FAB.
KM Amphibian skin.
SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

```

```

Query Match 25.7%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FY 2
DB 3 FF 4

RESULT 5
ID P82072 PRELIMINARY; PRT; 5 AA.
AC P82072.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELLIDIN 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCB1_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIBIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MM=655; METHOD=FAB.
KM Amphibian skin; Amidation.
FT MOD_RES 5
FT SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

```

```

Query Match 25.7%; Score 9; DB 13; Length 5;

```


Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
1:
3 FF 4

Db

RESULT 6
P82073 PRELIMINARY; PRT; 5 AA.
AC P82073;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE RUBELLIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN SECRETION;
RA Wapnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electrica. Comparison with the skin peptides from Litoria
rubella";
RU Aust J. Chem. 52:0-0(1989).
CC -I- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
ANTIBIOTIC ACTIVITY.
CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SQ SEQUENCE 5 AA; 570 MW; 71A9C9C862A0000 CRC64;

Query Match 25.7%; Score 9; DB 13; Length 5;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FY 2
1:
3 FF 4

Db

RESULT 7
Q9YQ10 PRELIMINARY; PRT; 7 AA.
AC Q9YQ10;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE HYPOHETICAL FUSION PROTEIN.
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99099045; Pubmed=9882359;
RA Izeta A., Smerdou C., Alonso S., Penzes Z., Mendez A., Piana-Duran J.,
RA Enjuanes L.;
RT "Replication and packaging of transmissible gastroenteritis
coronavirus-derived synthetic minigenomes";
RL J. Virol. 73:1535-1545(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95159435; Pubmed=7856095;
RA Elouet J., Raschaert D., Lambert P., Levy L., Yende P., Laude H.;
RT "Complete sequence (20 kilobases) of the polyprotein encoding gene 1
of transmissible gastroenteric virus";
RL Virology 206:817-822(1995).
RN [3]

RP SEQUENCE FROM N.A.
RX MEDLINE=88078100; Pubmed=2825819;
RA Raschaert D., Gelfi J., Laude H.;
RT "Enteric coronavirus TGEV: partial sequence of the genomic RNA its
RT organization and expression";
RL Biochimie 69:591-600(1987).
DR EMBL, AJ011462; CAA09625.1;
SQ SEQUENCE 7 AA; 927 MW; 69D6D7273B5726F0 CRC64;

Query Match 25.7%; Score 9; DB 12; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 YQL 4
1:
3 YLL 5

Db

RESULT 8
P82065 PRELIMINARY; PRT; 7 AA.
AC P82065;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-MAY-2000 (TREMblrel. 13, last annotation update)
DE TRYPTOPHYLLIN 5.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN SECRETION;
RA Steinhorn S.T., Wapnitz P.A., Waugh R.J., Bowle J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
Litoria rubella. The skin peptide profile as a probe for the study
of evolutionary trends of amphibians";
RU Aust. J. Chem. 49:955-963(1996).
CC -I- FUNCTION: MAY ACT AS A NEUROMODULATOR OR NEUOTRASMITTER.
CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -I- MASS SPECTROMETRY: MW=965; METHOD=FA-
KW Amphibian skin; Amidation; Neuropeptide.
FT MOD_RES 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 7 AMIDATION.
SQ SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 25.7%; Score 9; DB 13; Length 7;
Best Local Similarity 33.3%; Pred. No. 5.6e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQ 3
1:
5 FHR 7

Db

RESULT 9
P82101 PRELIMINARY; PRT; 7 AA.
AC P82101;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-MAY-2000 (TREMblrel. 13, last annotation update)
DE ELECTRIN 5.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OX NCBI_TaxID=104895;
RN [1]

RP SEQUENCE.
RC TISSUE-SKIN SECRETION;
RA Mabiniz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
Litoria electrica. Comparison with the skin peptides from Litoria
rubella";
RL Aust. J. Chem. 52:0-0(1999).
RW Amphibian skin; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 834 MW; 6DD05B076B0B5030 CRC64;

Query Match 25.7%; Score 9; DB 13; Length 7;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YQ 3
DB 2 YE 3

RESULT 10
OQ95945 PRELIMINARY; PRT; 7 AA.
AC OQ95945;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INSIDE INTRON 5 (FRAGMENT).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DD273-108;
RX MEDLINE=81069885; PubMed=6254986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
nucleotide sequence of the gene coding for subunit 1 of yeast
cytochrome oxidase";
RL J. Biol. Chem. 255:11927-11941(1980).
DR EMBL: V00694; CAA24066.1; -.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 7 AA; 859 MW; 75B7232362DC460 CRC64;

Query Match 22.9%; Score 8; DB 8; Length 7;
Best Local Similarity 25.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLA 5
DB 4 WKLS 7

RESULT 11
OQ9CSB3 PRELIMINARY; PRT; 7 AA.
AC OQ9CSB3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 0.7 KDA PROTEIN (FRAGMENT).
CN DIDI 10A-2B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=ROOT;
RX MEDLINE=21171025; PubMed=11277426;
RA Vercauteren I., Van Der Schueren E., Van Montagu M., Cheysen G.;
RT "Arabidopsis thaliana genes expressed in the early compatible
interaction with root-knot nematodes";
RL Mol. Plant Microbe Interact. 14:288-299(2001).
DR EMBL: AJ286350; CAB71014.2; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 7 AA; 719 MW; 6732C7287EB325D0 CRC64;

Query Match 22.9%; Score 8; DB 10; Length 7;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 LAL 6
DB 5 LVL 7

RESULT 12
OQ63668 PRELIMINARY; PRT; 7 AA.
AC OQ63668;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF2 PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DRAWLEY; TISSUE=KIDNEY;
RX MEDLINE=95396550; PubMed=767072;
RA Mandon B., Bellanger A.C., Elalouf J.M.;
RT "Inverse-PCR-mediated cloning of the promoter for the rat vasopressin
V2 receptor gene";
RL Pfleger's Arch. 430:12-18(1995).
DR EMBL: X83264; CAA58237.1; -.
SQ SEQUENCE 7 AA; 703 MW; 75A767287DC6D6F0 CRC64;

Query Match 22.9%; Score 8; DB 11; Length 7;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 QAL 6
DB 2 QAGL 5

RESULT 13
OQ6113 PRELIMINARY; PRT; 7 AA.
AC OQ6113;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE C-TERMINUS OF THE VIRAL REPLICASE (FRAGMENT).
OS Cherry leaf roll virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
OC Nepovirus.
OX NCBI_TaxID=12615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALNUT;
RA Botja M.;
RL Thesis (1992), Biologia Molecular y Virologia Vegetal, CIT-INIA.

RN [2] SEQUENCE FROM N.A.
 RP STRAIN=ALANT;
 RX MEDLINE=96124520; PubMed=8560786;
 RA Borja M., Sanchez F., Rowhani A., Bruening G., Ponz F.;
 RT "Long, nearly identical untranslated sequences at the 3' terminal
 RT regions of the genomic RNAs of cherry leafroll virus (walnut
 RT strain)." ;
 RL Virus Genes 10:245-252(1995).
 DR EMBL, Z34265; CAA84019.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 800 MW; 7417672EBDC60740 CRC64;

Query Match 22.9%; Score 8; DB 12; Length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 QUAL 6
 Db 2 QACL 5

RESULT 14
 P82100 PRELIMINARY; PRT; 5 AA.
 AC P82100;
 DT 01-MAY-2000 (Tremblrel, 13, Created)
 DT 02-MAY-2000 (Tremblrel, 13, Last sequence update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN SECRETION.
 RA Wadnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria electrica. Comparison with the skin peptides from Litoria
 RT rubella." ;
 RL Aust. J. Chem. 52:0-0(1999).
 KW Amphibian skin; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 20.0%; Score 7; DB 13; Length 5;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 LF 7
 Db 2 IT 3

RESULT 15
 P82181 PRELIMINARY; PRT; 6 AA.
 AC P82181;
 DT 01-JUN-2000 (Tremblrel, 14, Created)
 DT 01-JUN-2000 (Tremblrel, 14, Last sequence update)
 DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
 OS Spinacia oleracea (Spinach).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.

RC STRAIN=CV ALMARO; TISSUE=LEAF;
 RX MEDLINE=20435798; PubMed=10874046;
 RA Yamaguchi K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RT the 50 S subunit of an organelle ribosome (chloroplast)." ;
 RL J. Biol. Chem. 275:28466-28482(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR001790; Ribosomal_L10.
 DR InterPro: IPR002363; Ribosomal_L10_eub.
 DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
 DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 20.0%; Score 7; DB 10; Length 6;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 ALT 7
 Db 1 AIS 3

Search completed: July 15, 2002, 13:37:39
 Job time: 729 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:25:25 ; Search time 107.71 Seconds
(without alignments)
7.219 Million cell updates/sec

Title: US-09-712-819A-3
Perfect score: 35
Sequence: 1 FYDLALR 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 52936

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.1	6	19	AAV20207	Human beta-amyloid
2	54.3	5	18	AAW12505	Interleukin-6 anta
3	51.4	4	15	AAW46803	Phytase derived pe
4	51.4	5	18	AAW12486	Interleukin-6 anta
5	51.4	5	18	AAW11532	Interleukin-6 anta
6	51.4	6	13	AAW24623	Immunomodulatory p
7	51.4	6	18	AAW11531	Interleukin-6 anta
8	51.4	6	18	AAW11534	Interleukin-6 anta
9	51.4	6	19	AAW83888	Peptide specific a
10	51.4	6	19	AAW83891	Peptide specific a
11	51.4	6	19	AAW83892	Peptide specific a

12	18	51.4	6	19	AAW83872	Peptide specific a
13	18	51.4	6	19	AAW33781	Peptide #2 used in
14	18	51.4	6	19	AAW47263	Immunomodulatory p
15	18	51.4	6	22	AAW87728	Human Interleukin-
16	18	51.4	7	18	AAW11535	Interleukin-6 anta
17	18	51.4	7	18	AAW11530	Interleukin-6 anta
18	18	51.4	7	21	AAW10113	Human angiotensin-
19	18	51.4	7	22	AAW05547	Synthetic heptapep
20	18	51.4	7	22	AAW88727	Human Interleukin-
21	18	51.4	7	22	AAW72813	Antibacterial pept
22	17	48.6	5	18	AAW12502	Interleukin-6 anta
23	17	48.6	5	18	AAW12503	Interleukin-6 anta
24	17	48.6	5	18	AAW12504	Interleukin-6 anta
25	17	48.6	5	18	AAW12506	Interleukin-6 anta
26	17	48.6	5	18	AAW12497	Interleukin-6 anta
27	17	48.6	5	22	AAW51304	Anti-HIV peptide w
28	17	48.6	5	22	AAW51326	Anti-HIV peptide w
29	17	48.6	5	22	AAW51329	Anti-HIV peptide w
30	17	48.6	5	22	AAW51365	Anti-HIV peptide w
31	17	48.6	5	15	AAW65984	Peptide which bind
32	17	48.6	6	15	AAW65841	Anticoagulant nucle
33	17	48.6	6	19	AAW83854	Peptide specific a
34	17	48.6	6	19	AAW83726	Peptide specific a
35	17	48.6	6	19	AAW83603	Peptide specific a
36	17	48.6	7	15	AAW66001	Peptide which bind
37	17	48.6	7	15	AAW65858	Anticoagulant nucle
38	17	48.6	7	19	AAW46372	Peptide sequence f
39	17	48.6	7	22	AAW00247	Peptide fragment #
40	16	45.7	5	16	AAW89808	Melanotrophic rele
41	16	45.7	5	16	AAW89809	Melanotrophic rele
42	16	45.7	5	16	AAW89810	Melanotrophic rele
43	16	45.7	5	16	AAW89812	Melanotrophic rele
44	16	45.7	5	18	AAW12499	Interleukin-6 anta
45	16	45.7	5	18	AAW12500	Interleukin-6 anta

ALIGNMENTS

RESULT 1
AAV20207
ID AAV20207 standard; Protein; 6 AA.
XX
XX AAV20207;
DT 22-JUL-1999 (first entry)
XX
DE Human beta-amyloid precursor protein mutant fragment 32.
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumor antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HSP-71;
KW bel-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX Synthetic.
OS Homo sapiens.
OS
XX W09845322-A2.
XX
XX 15-OCT-1998.
PD
XX
XX 02-APR-1998; 98WO-IB00705.
XX
XX 10-APR-1997; 97US-0043163.
PR
XX
XX (UYUT-) RIKSUNIV UPRRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX WPI: 1998-609901/51.
 DR N-PSDB: AAX75753.

PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

PS Disclosure: Figure 2; 258bp; English.

XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGPC) and neuroendocrine specific protein A.

SQ Sequence 6 AA:

Query Match 57.1%; Score 20; DB 19; Length 6;
 Best Local Similarity 100.0%; Pred. No. 6.4e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLA 5
 ||||
 Db 3 yola 6

RESULT 2

AAW12505
 ID AAW12505 standard; peptide; 5 AA.

AC AAW12505;

DT 22-APR-1997 (first entry)

DE Interleukin-6 antagonist 70.

KW Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;
 KW skin; intestine; systemic lupus erythematosus; chronic rheumatism.

OS Synthetic.

FT Key Location/Qualifiers
 FT Modified-site 5 /note="amdated"

FN JP08311098-A.

PD 26-NOV-1996.

PF 22-MAY-1995; 95JP-0146742.

PR 22-MAY-1995; 95JP-0146742.

XX 5

PA (DAIL) DAICEL CHEM IND LTD.
 PA (FUJI) FUJISAWA PHARM CO LTD.
 XX WPI: 1997-061811/06.

PT Interleukin-6 antagonistic peptide(s) comprising arginine - useful
 PT for treating autoimmune, renal, skin and intestinal diseases

XX Example 70; Page 12; 20pp; Japanese.

CC The present peptide is a specific example of new interleukin-6
 CC antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
 CC acids or an amino group protecting group; Y is 1-5 amino acids, a
 CC carboxyl group protecting group or an amide; A is preferably Arg
 CC having an opt. protected guanidino group but can be any amino acid;
 CC D is Arg having an opt. protected guanidino group and B is preferably
 CC a Leu residue but can be any amino acid, including non-natural
 CC amino acids, opt. having a protected side-chain. The peptides are
 CC useful for treating autoimmune diseases (e.g. systemic lupus
 CC erythematosus or chronic rheumatism), renal, skin and intestinal
 CC diseases.

SQ Sequence 5 AA:

Query Match 54.3%; Score 19; DB 18; Length 5;
 Best Local Similarity 75.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOL 4
 ||:|
 Db 1 fye1 4

RESULT 3

AAW46803
 ID AAW46803 standard; Protein; 4 AA.

AC AAW46803;

DT 19-AUG-1994 (first entry)

DE Phytase derived peptide 418(3 phy).

KW pH 2.5; acid phosphatase; Trichoderma; host: Aspergillus; phytic acid;
 KW phytate degrading enzyme; PDE; removal; inositol hexaphosphoric acid;
 KW plant; feed composition; filtration.

OS Synthetic.

FN WO9403612-A.

PD 17-FEB-1994.

PF 30-JUL-1993; 93WO-FI00310.

PR 31-JUL-1992; 92US-0923724.

PA (ALKO-) ALKO LTD.

PT Cantrell M, Fagerstrom RB, Miettinen-Oinonen ASK;
 PT Nevalainen HK, Palojelimo MT, Piddington C, Rambosek JA;
 PT Torkkeli TK, Turunen MK.

XX WPI: 1994-065700/08.

PT Compns. contg. phytate degrading enzymes - obtd. by expression
 PT of their genes in Trichoderma, used partic. for producing animal
 PT feed compns.

XX Example 4; Page 43; 142pp; English.

CC The sequences given in AWA6793-824 are peptides derived from the

CC phytase protein. The phytase protein may be used in the composition
CC of the invention. The DNA encoding the phytase protein may be
CC introduced into a Trichoderma host which then expresses it and the
CC protein is collected from the culture medium. By using Trichoderma as
CC a host for Aspergillus phytase degrading enzymes such as this, a
CC totally different enzyme composition compared to that secreted from
CC Aspergillus results. The enzyme composition can be used for removal
CC of phytic acid or inositol hexaphosphoric acid from raw material,
CC particularly plant material. The composition is used in feed
CC compositions for animals. By using Trichoderma as a source of a
CC composition containing phytase degrading enzymes some difficult
CC downstream processing problems, eg. filtration, that occur with
CC similar Aspergillus compositions are avoided and yields are improved.

XX
SQ Sequence 4 AA;

Query Match 51.4%; Score 18; DB 15; Length 4;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQ 3
111
Db 1 fyt 3

RESULT 4

AAW12496
ID AAW12496 standard; peptide: 5 AA.

XX
AC AAW12496;

XX
DT 22-APR-1997 (first entry)

XX
DE Interleukin-6 antagonist 61.

XX
KM Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;

XX
KM skin; intestine; systemic lupus erythematosus; chronic rheumatism.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers

XX
FT Modified-site 5 /note= "amidated"

XX
PN JP08311098-A.

XX
PD 26-NOV-1996.

XX
PE 22-MAY-1995; 95JP-0146742.

XX
PR 22-MAY-1995; 95JP-0146742.

XX
PA (DAIL) DAICEL CHEM IND LTD.

XX
PA (FUJI) FUJISAWA PHARM CO LTD.

XX
DR WPI; 1997-061811/06.

XX
PT Interleukin-6 antagonistic peptide(s) comprising arginine - useful

XX
PT for treating autoimmune, renal, skin and intestinal diseases

XX
PS Example 61; Page 12; 20pp; Japanese.

XX
CC The present peptide is a specific example of new interleukin-6
CC antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
CC acids or an amino group protecting group; Y is 1-5 amino acids, a
CC carboxyl group protecting group or an amide; A is preferably Arg
CC having an opt. protected guanidino group but can be any amino acid;
CC D is Arg having an opt. protected guanidino group and B is preferably
CC a Leu residue but can be any amino acid, including non-natural
CC amino acids, opt. having a protected side-chain. The peptides are
CC useful for treating autoimmune diseases (e.g. systemic lupus
CC erythematosus or chronic rheumatism), renal, skin and intestinal

CC diseases.
XX
SQ Sequence 5 AA;

Query Match 51.4%; Score 18; DB 18; Length 5;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOL 4
111
Db 1 fyt 4

RESULT 5

AAW11532
ID AAW11532 standard; peptide: 5 AA.

XX
AC AAW11532;

XX
DT 22-APR-1997 (first entry)

XX
DE Interleukin-6 antagonist 5.

XX
KM Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;

XX
KM skin; intestine; systemic lupus erythematosus; chronic rheumatism.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers

XX
FT Modified-site 5 /note= "optionally amidated"

XX
PN JP08311098-A.

XX
PD 26-NOV-1996.

XX
PE 22-MAY-1995; 95JP-0146742.

XX
PR 22-MAY-1995; 95JP-0146742.

XX
PA (DAIL) DAICEL CHEM IND LTD.

XX
PA (FUJI) FUJISAWA PHARM CO LTD.

XX
DR WPI; 1997-061811/06.

XX
PT Interleukin-6 antagonistic peptide(s) comprising arginine - useful

XX
PT for treating autoimmune, renal, skin and intestinal diseases

XX
PS Example 5 and Example 6; Page 11; 20pp; Japanese.

XX
CC The present sequence covers two specific examples of new interleukin-6
CC antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
CC acids or an amino group protecting group; Y is 1-5 amino acids, a
CC carboxyl group protecting group or an amide; A is preferably Arg
CC having an opt. protected guanidino group but can be any amino acid;
CC D is Arg having an opt. protected guanidino group and B is preferably
CC a Leu residue but can be any amino acid, including non-natural
CC amino acids, opt. having a protected side-chain. The peptides are
CC useful for treating autoimmune diseases (e.g. systemic lupus
CC erythematosus or chronic rheumatism), renal, skin and intestinal
CC diseases.

XX
SQ Sequence 5 AA;

Query Match 51.4%; Score 18; DB 18; Length 5;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOL 4
111
Db 1 fyt 4

```

RESULT 6
AAR24623
ID AAR24623 standard; peptide: 6 AA.
XX
XX AAR24623;
AC
XX 03-DEC-1992 (first entry)
DT
XX Immunomodulatory peptide.
DE
XX Immunodeficiencies; immunosuppression; T-cell subset; immunotherapy;
KM inflammation; wounds; lymphocyte; vaccine.
XX
XX Synthetic.
OS
XX WO9209628-A.
XX
XX 11-JUN-1992.
PD
XX
XX 22-NOV-1991; 91WO-US08795.
PF
XX 23-NOV-1990; 90US-0617494.
PR
XX (IMMU-) IMMUNODYNAMICS INC.
PA
XX
XX Atkin A;
PI
XX WPI: 1992-217021/26.
DR
XX
XX New synthetic immunomodulatory peptide(s) - for treating
PT immunodeficiencies, immunosuppression and T-cell subset
PT deviations and immunotherapy of infections, inflammation, wounds
PT etc.
XX
XX Claim 10; Page 36; 52pp; English.
PS
XX
XX The immunomodulatory peptide is a specific example of a peptide cpd.
CC (or an acid or base salt) constructed by combination and/or
CC overlapping of the amino acid sequences A1B1XB2A2, A3B3XA4B4,
CC B5A5XA6B6, B7A7XB8A8, A9B9, A10A11, B10A12, (X= Ala, Gly,
CC Ile, Leu, Phe or Val, A1-A12 each- Arg, Asn, Gln, Lys, Phe or Val;
CC B1-B12 each- Asp, Glu, Tyr, Phe or Val. The synthetic peptide may
CC be used for immunomodulation of various immunodeficiencies and
CC immunosuppressed conditions, T-cell subset and lymphocyte deviations,
CC enhancement of a vaccine's efficacy, as well as for immunotherapy,
CC including infections, local or systemic complications of non-
CC infectious diseases, postoperative inflammations, wounds and burns.
CC See also AAR24583-R24701.
CC
XX
XX Sequence 6 AA:
SQ

```

```

Query Match 51.4%; Score 18; DB 13; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FYQL 4
   ||:|
DB 1 FYRL 4

```

```

KM Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;
KM skin; intestine; systemic lupus erythematosus; chronic rheumatism.
XX
XX Synthetic.
OS
XX key Location/Qualifiers
XX Modified-site 6 /note= "amdated"
XX
XX JP08311098-A.
XX
XX 26-NOV-1996.
PD
XX
XX 22-MAY-1995; 95JP-0146742.
PF
XX 22-MAY-1995; 95JP-0146742.
PR
XX 22-MAY-1995; 95JP-0146742.
XX
XX (DAIL ) DAICEL CHEM IND LTD.
XX (FUJI ) FUJISAWA PHARM CO LTD.
XX
XX WPI: 1997-061811/06.
DR
XX
XX Interleukin-6 antagonistic peptide(s) comprising arginine - useful
PT for treating autoimmune, renal, skin and intestinal diseases
PT
XX
XX Example 1; Page 11; 20pp; Japanese.
PS
XX
XX The present peptide is a specific example of new interleukin-6
CC antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
CC acids or an amino group protecting group; Y is 1-5 amino acids, a
CC carboxyl group protecting group or an amide; A is preferably Arg
CC having an opt. protected guanidino group but can be any amino acid;
CC D is Arg having an opt. protected guanidino group and B is preferably
CC a leu residue but can be any amino acid, including non-natural
CC amino acids, opt. having a protected side-chain. The peptides are
CC useful for treating autoimmune diseases (e.g. systemic lupus
CC erythematosus or chronic rheumatism), renal, skin and intestinal
CC diseases.
CC
XX
XX Sequence 6 AA:
SQ

```

```

Query Match 51.4%; Score 18; DB 18; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FYQL 4
   ||:|
DB 2 FYRL 5

```

```

RESULT 8
AAW11534
ID AAW11534 standard; peptide: 6 AA.
XX
XX AAW11534;
AC
XX 22-APR-1997 (first entry)
DT
XX Interleukin-6 antagonist 9.
DE
XX Interleukin-6 antagonist 9.
XX
XX Interleukin-6; IL-6; antagonist; inhibitor; autoimmune disease;
KM skin; intestine; systemic lupus erythematosus; chronic rheumatism.
XX
XX Synthetic.
OS
XX key Location/Qualifiers
XX Modified-site 6 /note= "amdated"
XX
XX JP08311098-A.
XX
XX 26-NOV-1996.
PD

```


XX 22-MAY-1995; 95JP-0146742.
XX
XX 22-MAY-1995; 95JP-0146742.
XX
PA (DAIL) DAICEL CHEM IND LTD.
PA (FUJI) FUJISAWA PHARM CO LTD.
XX
XX WPI; 1997-061811/06.
XX
XX Interleukin-6 antagonistic peptide(s) comprising arginine - useful
PT for treating autoimmune, renal, skin and intestinal diseases
XX
XX
PS Example 9; Page 11; 20pp; Japanese.

CC The present peptide is a specific example of new interleukin-6
CC antagonists of the general formula X-A-B-D-Y, where X is 1-10 amino
CC acids or an amino group protecting group; Y is 1-5 amino acids, a
CC carboxyl group protecting group or an amide; A is preferably Arg
CC having an opt. protected guanidino group but can be any amino acid;
CC D is Arg having an opt. protected guanidino group and B is preferably
CC a leu residue but can be any amino acid including non-natural
CC amino acids, opt. having a protected side-chain. The peptides are
CC useful for treating autoimmune diseases (e.g. systemic lupus
CC erythematosus or chronic rheumatism), renal, skin and intestinal
CC diseases.
XX
XX

SQ Sequence 6 AA;

Query Match 51.4%; Score 18; DB 18; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQL 4
1111
Db 1 fyt 4

RESULT 9

AAW83888 standard; peptide; 6 AA.

AC AAW83888;

DT 14-JAN-1999 (first entry)

DE Peptide specific antigen (PSA) cleaved substrate sequence.

KW Human prostate specific antigen; PSA; substrate; prostate cancer;
KW metastatic prostate cancer; PSA inhibitor.

OS Synthetic.

PN WO9840738-A1.

PD 17-SEP-1998.

PF 09-MAR-1998; 98WO-US04557.

PR 11-MAR-1997; 97US-0040174.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PI Blobel C, Roghani M, Rothman J;

XX WPI; 1998-520833/44.

XX Improved human prostate specific antigen substrates - useful to,
PT e.g. measure prostate specific antigen activity, identify
PT therapeutically useful inhibitors or diagnose metastatic prostate
PT cancer

PS Claim 17; Fig 18; 147pp; English.

XX AAW83856-94 represent peptides cleaved by 50 microgram/ml human prostate
CC specific antigen (PSA). They correspond to the consensus sequence
CC AAW83855. The PSA substrates are useful in assays which measure PSA
CC activity. Such assays may provide a clinical evaluation of patients
CC suffering from, or at risk for, prostate cancer. Elevated levels of
CC PSA in serum are widely used as a marker of prostate cancer, and the
CC substrates can be used e.g. to measure the amount of active PSA in
CC patient serum to diagnose, or determine the malignant potential of,
CC a patient's prostate disease. They are especially useful to diagnose
CC metastatic prostate cancer. The substrates can also be used in assays
CC to identify PSA inhibitors, which can be administered to treat or
CC prevent prostate cancer.

SQ Sequence 6 AA;

Query Match 51.4%; Score 18; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQ 3
1111
Db 1 fyt 3

RESULT 10

AAW83891 standard; peptide; 6 AA.

AC AAW83891;

DT 14-JAN-1999 (first entry)

DE Peptide specific antigen (PSA) cleaved substrate sequence.

KW Human prostate specific antigen; PSA; substrate; prostate cancer;
KW metastatic prostate cancer; PSA inhibitor.

OS Synthetic.

PN WO9840738-A1.

PD 17-SEP-1998.

PF 09-MAR-1998; 98WO-US04557.

PR 11-MAR-1997; 97US-0040174.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PI Blobel C, Roghani M, Rothman J;

XX WPI; 1998-520833/44.

XX Improved human prostate specific antigen substrates - useful to,
PT e.g. measure prostate specific antigen activity, identify
PT therapeutically useful inhibitors or diagnose metastatic prostate
PT cancer

PS Claim 17; Fig 18; 147pp; English.

XX AAW83856-94 represent peptides cleaved by 50 microgram/ml human prostate
CC specific antigen (PSA). They correspond to the consensus sequence
CC AAW83855. The PSA substrates are useful in assays which measure PSA
CC activity. Such assays may provide a clinical evaluation of patients
CC suffering from, or at risk for, prostate cancer. Elevated levels of
CC PSA in serum are widely used as a marker of prostate cancer, and the
CC substrates can be used e.g. to measure the amount of active PSA in
CC patient serum to diagnose, or determine the malignant potential of,
CC a patient's prostate disease. They are especially useful to diagnose
CC metastatic prostate cancer. The substrates can also be used in assays

CC to identify PSA inhibitors, which can be administered to treat or
CC prevent prostate cancer.
XX
SQ Sequence 6 AA;

Query Match 51.4%; Score 18; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYO 3
|||
Db 1 fYq 3

RESULT 11
AAW83892
ID AAW83892 standard; peptide: 6 AA.
XX
AC AAW83892;
XX
DE 14-JAN-1999 (first entry)
XX
DE Peptide specific antigen (PSA) cleaved substrate sequence.
XX
KW Human prostate specific antigen; PSA; substrate; prostate cancer;
XX metastatic prostate cancer; PSA inhibitor.
XX
OS Synthetic.
XX
PN WO9840738-A1.
XX
PD 17-SEP-1998.
XX
PF 09-MAR-1998; 98WO-US04557.
XX
PR 11-MAR-1997; 97US-0040174.
XX
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Blobel C, Roghani M, Rothman J;
XX
DR WPI; 1998-520833/44.
XX
PT Improved human prostate specific antigen substrates - useful to,
PT e.g. measure prostate specific antigen activity; identify
PT therapeutically useful inhibitors or diagnose metastatic prostate
PT cancer.
PS Claim 17; Fig 18; 147pp; English.
XX
CC AAW83856-94 represent peptides cleaved by 50 microgram/ml human prostate
CC specific antigen (PSA). They correspond to the consensus sequence
CC AAW83855. The PSA substrates are useful in assays which measure PSA
CC activity. Such assays may provide a clinical evaluation of patients
CC suffering from, or at risk for, prostate cancer. Elevated levels of
CC PSA in serum are widely used as a marker of prostate cancer, and the
CC substrates can be used e.g. to measure the amount of active PSA in
CC patient serum to diagnose, or determine the malignant potential of,
CC a patient's prostate disease. They are especially useful to diagnose
CC metastatic prostate cancer. The substrates can also be used in assays
CC to identify PSA inhibitors, which can be administered to treat or
CC prevent prostate cancer.
SQ Sequence 6 AA;

Query Match 51.4%; Score 18; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYO 3
|||,

Db 1 fYq 3

RESULT 12
AAW83872
ID AAW83872 standard; peptide: 6 AA.
XX
AC AAW83872;
XX
DE 14-JAN-1999 (first entry)
XX

DE Peptide specific antigen (PSA) cleaved substrate sequence.
XX
KW Human prostate specific antigen; PSA; substrate; prostate cancer;
XX metastatic prostate cancer; PSA inhibitor.
XX

OS Synthetic.

PN WO9840738-A1.

PD 17-SEP-1998.

PF 09-MAR-1998; 98WO-US04557.

PR 11-MAR-1997; 97US-0040174.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PI Blobel C, Roghani M, Rothman J;

DR WPI; 1998-520833/44.

PT Improved human prostate specific antigen substrates - useful to,
PT e.g. measure prostate specific antigen activity; identify
PT therapeutically useful inhibitors or diagnose metastatic prostate
PT cancer

PS Claim 17; Fig 18; 147pp; English.

XX
CC AAW83856-94 represent peptides cleaved by 50 microgram/ml human prostate
CC specific antigen (PSA). They correspond to the consensus sequence
CC AAW83855. The PSA substrates are useful in assays which measure PSA
CC activity. Such assays may provide a clinical evaluation of patients
CC suffering from, or at risk for, prostate cancer. Elevated levels of
CC PSA in serum are widely used as a marker of prostate cancer, and the
CC substrates can be used e.g. to measure the amount of active PSA in
CC patient serum to diagnose, or determine the malignant potential of,
CC a patient's prostate disease. They are especially useful to diagnose
CC metastatic prostate cancer. The substrates can also be used in assays
CC to identify PSA inhibitors, which can be administered to treat or
CC prevent prostate cancer.
SQ Sequence 6 AA;

Query Match 51.4%; Score 18; DB 19; Length 6;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYO 3
|||
Db 1 fYq 3

RESULT 13

AAW33781
ID AAW33781 standard; peptide: 6 AA.

AC AAW33781;

DE 19-JUN-1998 (first entry)

DE Peptide #2 used in immunomodulating dimer peptide.

XX Immunomodulating dimer; immunosuppressant drug; CTL activation;
 KW transplantation; autoimmune disease; Class I HLA-B alpha-1 domain;
 KW rejection.
 XX Synthetic.
 OS Homo sapiens.
 PN WO9744351-A1.
 PD 27-NOV-1997.
 XX 22-MAY-1997; 97WO-US08689.
 XX 24-MAY-1996; 96US-0653294.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Beulow R, Clayberger C, Krensky AM;
 DR WPI: 1998-086530/08.
 XX New immunomodulating dimer peptide(s) - based on a Class I HLA-B
 PT alpha-1 domain, used for preventing rejection of transplants or
 PT treating autoimmune diseases
 XX
 PS Claim 15; Page 35; 41pp; English.
 CC This sequence represents a specifically claimed peptide which forms part
 CC of the immunomodulating dimer peptides of the invention. A peptide-type
 CC compound or variant is claimed which has immunomodulating activity,
 CC including the N-terminal acylated and/or C-terminal amidated or
 CC esterified forms of up to 60 amino acids, where the peptide-type compound
 CC comprises the formula: A-B, where A, B = (R aa76-77), (aa79-84) or
 CC (aa84-79) (Laa77-76R); aa76 = E or V; aa77 = D, S or N; aa79 = R or G;
 CC aa80 = I or N; aa81, aa84 = a hydrophobic or small amino acid; aa82 = R
 CC or L; aa83 = G or R; and aa represents amino acid. The sequence in the
 CC brackets may optionally be absent or truncated at any peptide type bond
 CC within the brackets. The compounds comprise amino acid sequences related
 CC to a Class I HLA-B alpha domain (positions 79-84). They can be used to
 CC inhibit cytotoxic T-lymphocytes (CTL) from undesirably attacking cells in
 CC a host or in vitro. They can also be used in combination with antigenic
 CC peptides or proteins of interest to activate CTLs. They can also inhibit
 CC the proliferation of T cells in response to anti-CD3. The peptide can be
 CC used for preventing rejection of transplants or for treating autoimmune
 CC diseases, e.g. diabetes, rheumatoid arthritis and lupus erythematosus.
 CC The products can also be used for detection and diagnosis.
 XX
 SQ Sequence 6 AA:
 Query Match 51.4%; Score 18; DB 19; Length 6;
 Best Local Similarity 60.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 2 YQAL 6
 I::I:
 Db 1 yrlal 5
 RESULT 14
 AAW47263
 ID AAW47263 standard; peptide: 6 AA.
 XX
 AC AAW47263;
 XX
 DT 22-MAY-1998 (first entry)
 XX
 DE Immunomodulatory peptide.
 XX
 KW Immunomodulator; Class I HLA-B alpha-1 domain; inhibition;
 KW transplant rejection; treatment; autoimmune disease.
 XX

OS Homo sapiens.
 OS Synthetic.
 FH Key location/Qualifiers
 FT Misc-difference 1..6
 FT /note= "at least one of the amino acids is the
 D-isomer"
 PN WO9744052-A1.
 PD 27-NOV-1997.
 XX 23-APR-1997; 97WO-US06705.
 XX 22-MAY-1996; 96US-0651650.
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 PI Clayberger C, Krensky AM;
 DR WPI: 1998-018220/02.
 XX Novel immunomodulatory peptide-type compound - useful for inhibiting
 PT transplant rejection
 PT
 PS Claim 10; Page 36; 41pp; English.
 CC The present sequence is an immunomodulatory peptide, which
 CC comprises a Class I HLA-B alpha-1 domain sequence. It can be used
 CC in a pharmaceutical composition together with a subtherapeutic dose
 CC of an immunosuppressant, to extend the period of acceptance of a
 CC transplant from a major histocompatibility complex (MHC) unmatched
 CC donor, i.e. to inhibit transplant rejection. It can also be used in
 CC the treatment of autoimmune diseases.
 CC Peptides using the D-form amino acids are more effective
 CC immunomodulators than their diastereomers or enantiomers.
 XX
 SQ Sequence 6 AA:
 Query Match 51.4%; Score 18; DB 19; Length 6;
 Best Local Similarity 60.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 2 YQAL 6
 I::I:
 Db 1 yrlal 5
 RESULT 15
 AAB88728
 ID AAB88728 standard; Peptide: 6 AA.
 XX
 AC AAB88728;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Human interleukin-6 domain III fragment #16.
 XX
 KW Interleukin-6; ligand; IL-6 receptor; antagonist; cancer; allergy;
 KW rheumatoid arthritis; diabetes; multiple sclerosis; infection;
 KW autoimmune disease; inflammatory disease.
 XX
 OS Homo sapiens.
 OS
 PN WO200116166-A2.
 PD 08-MAR-2001.
 XX
 PF 25-AUG-2000; 2000WO-US23490.
 XX
 PR 27-AUG-1999; 99US-0151277.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Saxinger C;
XX

DR WPI: 2001-244395/25.

XX Polypeptides which inhibit the binding of Interleukin (IL)-6 ligand
PT with the IL-6 receptor, and the nucleic acids that encode them, useful
PT for treating e.g. inflammation and autoimmune diseases -
XX

PS Example 3: Page 46; 98pp; English.
XX

CC The present invention describes a number of peptides which are able to
CC bind to the human interleukin-6 (IL-6) receptor and inhibit IL-6 binding.
CC These are useful in the treatment of diseases associated with abnormal
CC IL-6 expression, including multiple myeloma, plasmacytoma, haematological
CC diseases such as plasma cell dyscrasias, leukaemia and lymphoma,
CC mesangial proliferative glomerulonephritis, polyclonal B cell activation
CC conditions, allergies, rheumatoid arthritis, diabetes, multiple
CC sclerosis, septic shock, infections, post-menopausal osteoporosis,
CC chronic immune deficiency, autoimmune diseases and inflammatory diseases.
XX

SQ Sequence 6 AA;

Query Match

51.4%; Score 18; DB 22; Length 6;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOL 4
11:1
Db 3 Tyr1 6

Search completed: July 15, 2002, 13:25:26
Job time: 1457 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:26:13 : Search time 40.89 seconds
(without alignments)
4.181 Million cell updates/sec

Title: US-09-712-819A-3
Sequence: 1 FLYOLALT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 39160

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	60.0	7	4	US-09-183-217-7	Sequence 7, Appl
2	18	51.4	4	1	US-07-923-724-30	Sequence 30, Appl
3	18	51.4	4	2	US-08-609-425A-30	Sequence 30, Appl
4	18	51.4	4	2	US-08-374-652C-22	Sequence 22, Appl
5	4	18	4	2	US-08-350-884-30	Sequence 30, Appl
6	17	48.6	7	1	US-08-709-173-30	Sequence 30, Appl
7	17	48.6	7	2	US-08-709-173-30	Sequence 30, Appl
8	17	48.6	7	3	US-08-946-329A-26	Sequence 26, Appl
9	17	48.6	7	4	US-08-567-357A-26	Sequence 26, Appl
10	17	48.6	7	4	US-08-729-743A-26	Sequence 26, Appl
11	17	48.6	7	4	US-09-457-046B-40	Sequence 40, Appl
12	16	45.7	5	1	US-08-238-089-23	Sequence 23, Appl
13	16	45.7	5	1	US-08-238-089-25	Sequence 23, Appl
14	16	45.7	5	1	US-08-432-651A-23	Sequence 23, Appl
15	16	45.7	5	1	US-08-432-651A-23	Sequence 23, Appl
16	16	45.7	5	3	US-08-962-962C-23	Sequence 23, Appl
17	16	45.7	5	3	US-08-962-962C-25	Sequence 25, Appl
18	16	45.7	5	5	PCT-US95-05560-23	Sequence 23, Appl
19	16	45.7	5	5	PCT-US95-05560-25	Sequence 25, Appl
20	16	45.7	6	5	PCT-US93-01669-50	Sequence 50, Appl
21	16	45.7	7	1	US-08-193-977-15	Sequence 15, Appl
22	16	45.7	7	4	US-08-354-240A-8	Sequence 8, Appl
23	16	45.7	7	4	US-09-230-081-10	Sequence 10, Appl
24	16	45.7	7	4	US-09-258-754-271	Sequence 271, App
25	16	45.7	7	4	US-09-042-107-271	Sequence 271, App
26	16	45.7	7	4	US-09-457-046B-39	Sequence 39, App
27	16	45.7	7	6	5496721-5	Patent No. 5496721

28	15	42.9	4	6	5237466-26	Patent No. 5237466
29	15	42.9	5	1	US-08-392-646-12	Sequence 12, Appl
30	15	42.9	5	1	US-08-392-646-13	Sequence 13, Appl
31	15	42.9	5	2	US-08-612-785B-21	Sequence 21, Appl
32	15	42.9	5	3	US-08-469-141A-40	Sequence 40, Appl
33	15	42.9	5	4	US-08-703-675C-34	Sequence 34, Appl
34	15	42.9	5	4	US-08-617-267C-21	Sequence 21, Appl
35	15	42.9	5	4	US-08-617-267C-40	Sequence 40, Appl
36	15	42.9	5	5	PCT-US95-13794-40	Sequence 40, Appl
37	15	42.9	6	1	US-08-106-493A-3	Sequence 3, Appl
38	15	42.9	6	1	US-08-429-264-3	Sequence 3, Appl
39	15	42.9	6	1	US-08-374-843B-27	Sequence 27, Appl
40	15	42.9	6	2	US-08-317-310A-54	Sequence 54, Appl
41	15	42.9	6	2	US-08-905-420-27	Sequence 27, Appl
42	15	42.9	6	2	US-08-540-412-114	Sequence 114, Appl
43	15	42.9	6	2	US-08-540-412-115	Sequence 115, Appl
44	15	42.9	6	2	US-08-540-412-135	Sequence 135, Appl
45	15	42.9	6	2	US-08-540-412-159	Sequence 159, Appl

ALIGNMENTS

RESULT 1
US-09-183-217-7
Sequence 7, Application US/09183217A
Patent No. 6153194
GENERAL INFORMATION:
APPLICANT: Skare, Jonathan T.
APPLICANT: Shang, Ellen S.
APPLICANT: Champion, Cheryl I.
APPLICANT: Blanco, David R.
APPLICANT: Miller, James N.
APPLICANT: Lovett, Michael A.
APPLICANT: Mizabekov, Tajib A.
APPLICANT: Kagan, Bruce L.
APPLICANT: Tempst, Paul
APPLICANT: Foley, Denise M.
TITLE OF INVENTION: BORRELLIA BURGDORFERI OUTER MEMBRANE PROTEINS
FILE REFERENCE: UC Case No. 6153194 96-059-3/Skare et al.
CURRENT FILING DATE: 1998-10-29
PRIORITY FILING DATE: 1997-01-22
PRIORITY FILING DATE: 1997-01-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 7
LENGTH: 7
TYPE: PRT
ORGANISM: Borrelia burgdorferi
US-09-183-217-7

Query Match 60.0%; Score 21; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YOLALT 7
Db 1 YKLGIT 6
RESULT 2
US-07-923-724-30
Sequence 30, Application US/07923724
Patent No. 5780292
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Paloheimo, Marja T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.

APPLICANT: Ramdasek, John A.
APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes
TITLE OF INVENTION: In Trichoderma
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/923,724
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 1050.0240004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: both
US-07-923-724-30

Query Match 51.4%; Score 18; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQ 3
111
DB 1 FYQ 3

RESULT 3
US-08-609-426A-30
Sequence 30, Application US/08609426A
Patent No. 5830733
GENERAL INFORMATION:
APPLICANT: Nevalainen, Helena K.M.
APPLICANT: Palohelmo, Marja T.
APPLICANT: Miettinen-Oinonen, Arja S.K.
APPLICANT: Torkkeli, Tuula K.
APPLICANT: Cantrell, Michael
APPLICANT: Piddington, Christopher S.
APPLICANT: Ramdasek, John A.
APPLICANT: Turunen, Marja K.
APPLICANT: Fagerstr m, Richard B.
APPLICANT: Houston, Christine S.
TITLE OF INVENTION: Production of Phytase Degrading Enzymes
TITLE OF INVENTION: In Trichoderma
NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/609,426A
FILING DATE: 01-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,724
FILING DATE: 31-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/496,155
FILING DATE: 19-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/044,077
FILING DATE: 29-APR-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 8610600
FILING DATE: 30-APR-1986
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Grant E.
REGISTRATION NUMBER: P-41,264
REFERENCE/DOCKET NUMBER: 1050.0080001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2540
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
TOPOLOGY: both
US-08-609-426A-30

Query Match 51.4%; Score 18; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQ 3
111
DB 1 FYQ 3

RESULT 4
US-08-374-652C-22
Sequence 22, Application US/08374652C
Patent No. 5834286
GENERAL INFORMATION:
APPLICANT: NEVALAINEN, HELENA K.M.
APPLICANT: PALOHEIMO, MARJA T.
APPLICANT: FAGERSTROM, RICHARD B.
APPLICANT: MIETTINEN-OINONEN, ARJA S.
APPLICANT: TURUNEN, MARJA K.
APPLICANT: RAMDASEK, JOHN A.
APPLICANT: PIDDINGTON, CHRISTOPHER S.
APPLICANT: HOUSTON, CHRISTINE S.
APPLICANT: CANTRELL, MICHAEL A.
TITLE OF INVENTION: RECOMBINANT CELLS, DNA CONSTRUCTS,
TITLE OF INVENTION: VECTORS AND METHODS FOR EXPRESSING PHYTASE DEGRADING
TITLE OF INVENTION: ENZYMES IN DESIRED RATIOS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600

CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,652C
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07058
FILING DATE: 27-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/925,401
FILING DATE: 31-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: REED, GRANT E.
REGISTRATION NUMBER: 41,264
REFERENCE/DOCKET NUMBER: 1050.071001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-374-652C-22

Query Match 51.4%; Score 18; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYO 3
111
Db 1 FYO 3

RESULT 5
US-08-350-884-30
Sequence 30, Application US/08350884
Patent No. 5585258
GENERAL INFORMATION:
APPLICANT: HOUGHTON, MICHAEL
APPLICANT: CHOO, QUI LIM
APPLICANT: KUO, GEORGE
TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/350,884
FILING DATE: 06-DEC-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/680,296
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: CIOTTI, THOMAS E.
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 22300-20100.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-350-884-30

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 TOTAL 6
1:11
Db 2 YDIAL 6

RESULT 6
US-08-709-173-30
Sequence 30, Application US/08709173
Patent No. 5712145
GENERAL INFORMATION:
APPLICANT: HOUGHTON, MICHAEL
APPLICANT: CHOO, QUI LIM
APPLICANT: KUO, GEORGE
TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,173
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/680,296
FILING DATE: 04-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: CIOTTI, THOMAS E.
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 22300-20100.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-709-173-30

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YOTAL 6
| : | |
Db 2 YDIAL 6

RESULT 7
US-08-709-177-30

; Sequence 30, Application US/08709177
; Patent No. 5885799

; GENERAL INFORMATION:

; APPLICANT: HOGGTON, MICHAEL

; APPLICANT: CHOO, QUI LIM

; APPLICANT: KUO, GEORGE

; TITLE OF INVENTION: HEPATITIS C VIRUS PROTEASE

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORRISON & FOERSTER

; STREET: 755 Page Mill Road

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/709,177

; FILING DATE: 06-SEP-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/680,296

; FILING DATE: 04-APR-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: CIOTTE, THOMAS E.

; REGISTRATION NUMBER: 21,013

; REFERENCE/DOCKET NUMBER: 22300-20100,20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 813-5600

; TELEFAX: (415) 494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 30:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-709-177-30

Query Match 48.6%; Score 17; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 YOTAL 6
| : | |
Db 2 YDIAL 6

RESULT 8
US-08-946-329A-26

; Sequence 26, Application US/08946329A

; Patent No. 6057091

; GENERAL INFORMATION:

; APPLICANT: Beachy, Phillip A.

; APPLICANT: Porter, Jeffrey A.

; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

; NUMBER OF SEQUENCES: 109

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: Windows 95

; SOFTWARE: FASTSEQ for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/946,329A

; FILING DATE: 07-OCT-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/061,323

; FILING DATE: 07-OCT-1996

; APPLICATION NUMBER: 08/729,743

; FILING DATE: 10-JUL-1996

; APPLICATION NUMBER: 08/567,357

; FILING DATE: 04-DEC-1995

; APPLICATION NUMBER: 08/349,498

; FILING DATE: 02-DEC-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Hallie, Lisa A.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07265/140001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619/678-5070

; TELEFAX: 619/678-5099

; INFORMATION FOR SEQ ID NO: 26:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; US-08-946-329A-26

Query Match 48.6%; Score 17; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LALT 7
| | | |
Db 1 LALT 4

RESULT 9
US-08-567-357A-26

; Sequence 26, Application US/08567357A

; Patent No. 6132728

; GENERAL INFORMATION:

; APPLICANT: Beachy, Phillip A.

; APPLICANT: Moon, Randall T.

; APPLICANT: Porter, Jeffrey A.

; TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: Windows 95

; SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/567,357A
FILING DATE: 04-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/080001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-567-357A-26

Query Match 48.6%; Score 17; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LAIT 7
1111
DB 1 LAIT 4

RESULT 10
US-08-729-743A-26
Sequence 26, Application US/08729743A
Patent No. 6214794
GENERAL INFORMATION:
APPLICANT: Beachy, Philip A.
APPLICANT: Moon, Randall T.
APPLICANT: Porter, Jeffrey A.
TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,743A
FILING DATE: 07-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/567,357
FILING DATE: 04-DEC-1995
APPLICATION NUMBER: 08/349,498
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07265/099001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-729-743A-26

Query Match 48.6%; Score 17; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LAIT 7
1111
DB 1 LAIT 4

RESULT 11
US-09-457-046B-40
Sequence 40, Application US/09457046B
Patent No. 6287835
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney et al.
TITLE OF INVENTION: Transacylases of the Pacitaxel Biosynthetic Pathway
FILE REFERENCE: 53679
CURRENT APPLICATION NUMBER: US/09/457,046B
CURRENT FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-457-046B-40

Query Match 48.6%; Score 17; DB 4; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLA 5
1111
DB 1 YYPPLA 5

RESULT 12
US-08-238-089-23
Sequence 23, Application US/08238089
Patent No. 5589460
GENERAL INFORMATION:
APPLICANT: Abajian, Henry B.
APPLICANT: Hlavka, Joseph J.
TITLE OF INVENTION: 3HT-, 4HT-, 5HT-, 6HT-, AND POLYPEPTIDES
TITLE OF INVENTION: AND THEIR THERAPEUTIC USE AS AN ANTIDEPRESSANT AGENT
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,089
FILING DATE: 04-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Jr., Lewis F.

REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 3297-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /label=Gly-NH2
OTHER INFORMATION: /note="A modified Gly residue: an amine group
OTHER INFORMATION: replaces a hydroxyl group at the carboxy terminus."
US-08-238-089-23

Query Match 45.7%; Score 16; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FYOL 4
Db 1 FYPL 4

RESULT 13
US-08-238-089-25
Sequence 25, Application US/08238089
Patent No. 5589460
GENERAL INFORMATION:
APPLICANT: Abajian, Henry B.
APPLICANT: No. 55894601e, John F.
APPLICANT: Hlavka, Joseph J.
TITLE OF INVENTION: TRI-, TETRA-, PENTA-, AND POLYPEPTIDES
TITLE OF INVENTION: AND THEIR THERAPEUTIC USE AS AN ANTIDEPRESSANT AGENT
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/238,089
FILING DATE: 04-MAY-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Jr., Lewis F.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 3297-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site

LOCATION: 3
OTHER INFORMATION: /label=4Hyp
OTHER INFORMATION: /note="Amino acid #3 is either cis- or trans- 4Hyp"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /label=Gly-NH2
OTHER INFORMATION: /note="A modified Gly residue: an amine group
OTHER INFORMATION: replaces a hydroxyl group at the carboxy terminus."
US-08-238-089-25

Query Match 45.7%; Score 16; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FYOL 4
Db 1 FYXL 4

RESULT 14
US-08-432-651A-23
Sequence 23, Application US/08432651A
Patent No. 5767083
GENERAL INFORMATION:
APPLICANT: Abajian, Henry B.
APPLICANT: No. 57670831e, John F.
APPLICANT: Hlavka, Joseph J.
TITLE OF INVENTION: TRI-, TETRA-, PENTA-, AND POLYPEPTIDES
TITLE OF INVENTION: AND THEIR THERAPEUTIC USE AS AN ANTIDEPRESSANT AGENT
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,651A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Jr., Lewis F.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 3297-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /label=Gly-NH2
OTHER INFORMATION: /note="A modified Gly residue: an amine group
OTHER INFORMATION: replaces a hydroxyl group at the carboxy terminus."
US-08-432-651A-23

Query Match 45.7%; Score 16; DB 1; Length 5;

Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOL 4
111
Db 1 FYPL 4

Search completed: July 15, 2002, 13:26:14
Job time: 494 sec

RESULT 15
US-08-432-651A-25
Sequence 25, Application US/08432651A
Patent No. 5767083
GENERAL INFORMATION:
APPLICANT: Abajian, Henry B.
APPLICANT: No. 57670831e, John F.
APPLICANT: Hlavka, Joseph J.
TITLE OF INVENTION: TRI-, TETRA-, PENTA-, AND POLYPEPTIDES
TITLE OF INVENTION: AND THEIR THERAPEUTIC USE AS AN ANTIDEPRESSANT AGENT
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lewis F. Gould, Jr.
STREET: 1700 Market Street, Suite 3232
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,651A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Gould, Jr., Lewis F.
REGISTRATION NUMBER: 25,057
REFERENCE/DOCKET NUMBER: 3297-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 575-6020
TELEFAX: (215) 575-6015
INFORMATION FOR SEQ ID NO. 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /label-4HYP
OTHER INFORMATION: /note- "Amino acid #3 is either cis- or trans-
OTHER INFORMATION: 4HYP"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 5
OTHER INFORMATION: /label- Gly-NH2
OTHER INFORMATION: /note- "A modified Gly residue: an amine group
OTHER INFORMATION: replaces a hydroxyl group at the carboxy
OTHER INFORMATION: terminus."
US-08-432-651A-25

Query Match 45.7%; Score 16; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOL 4
111
Db 1 FYXL 4

10

S09652
 hypothetical protein (aac2 3' region) - Enterobacter cloacae (fragment)
 C:Species: Enterobacter cloacae
 C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 08-Oct-1999
 C:Accession: S09652
 R:Villegant, J.S.; Keteleer-van Galen, P.A.G.; van de Klundert, J.A.M.
 Antimicrob. Agents Chemother. 33, 1153-1159, 1989
 A:Title: Nucleotide sequence of the aac2 gene, a gentamicin resistance determinant invc
 A:Reference number: S09651; MUID:90024972
 A:Accession: S09652
 A:Molecule type: DNA
 A:Residues: 1-7 <VLIX>
 A:Cross-references: EMBL:X51534; NID:g40878; PIDN:CAA35914.1; PID:g581034

Query Match 34.8%; Score 16; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIW 7
 : 1 1
 DB 1 MIW 4

RESULT 4
 A28719
 thymic humoral factor gamma-2 - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 18-Jun-1993
 C:Accession: A28719
 R:Burbstein, Y.; Buchner, V.; Pecht, M.; Trainin, N.
 Biochemistry 27, 4086-4071, 1988
 A:Title: Thymic humoral factor gamma-2: purification and amino acid sequence of an immun
 A:Reference number: A28719; MUID:88326920
 A:Accession: A28719
 A:Molecule type: Protein
 A:Residues: 1-8 <BUR>

Query Match 34.8%; Score 16; DB 2; Length 8;
 Best Local Similarity 75.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PKL 4
 : 1 1
 DB 5 PKPL 8

RESULT 5
 G37196
 bradykinin-potentiating peptide 7 - island jararaca
 C:Species: Bothrops insularis (Island Jararaca)
 C>Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
 C:Accession: G37196
 R:Cintre, A.C.O.; Vieira, C.A.; Giglio, J.R.
 J. Protein Chem. 9, 221-227, 1990
 A:Title: Primary structure and biological activity of bradykinin potentiating peptides f
 A:Reference number: A37196; MUID:90351557
 A:Accession: G37196
 A>Status: Preliminary
 A:Molecule type: Protein
 A:Residues: 1-5 <CIN>
 A:Keywords: pyroglutamic acid
 F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 32.6%; Score 15; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WA 8
 : 1 1
 DB 3 WA 4

RESULT 6
 PD0028
 pen-kinin 2 - penaeid shrimp (Penaeus vannamei) (fragment)
 C:Species: Penaeus vannamei
 C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000
 C:Accession: PD0028
 R:Nieto, J.; Veleert, D.; Derna, R.; Waelkens, E.; Gerslaens, A.; Coast, G.; Devree
 Biochem. Biophys. Res. Commun. 248, 406-411, 1998
 A:Title: Identification of one tachykinin- and two kinin-related peptides in the brai
 A:Reference number: PD0027; MUID:98342103
 A:Accession: PD0028
 A:Molecule type: Protein
 A:Residues: 1-6 <NIE>
 A:Comment: This peptide belongs to myotropic neuropeptides.

Query Match 32.6%; Score 15; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WA 8
 : 1 1
 DB 5 WA 6

RESULT 7
 PD0727
 H2 class I protein [imported] - rice (fragment)
 C:Species: Oryza sativa (rice)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: PD0727
 R:Komatsu, S.; Kajiwara, H.; Hirano, H.
 Theor. Appl. Genet. 86, 935-942, 1993
 A:Title: A rice protein library: a data-file of rice proteins separated by two-dimens
 A:Reference number: PD0696
 A:Accession: PD0727
 A>Status: Preliminary
 A:Molecule type: Protein
 A:Residues: 1-7 <KOM>

Query Match 32.6%; Score 15; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WA 8
 : 1 1
 DB 1 WA 2

RESULT 8
 E48394
 glycoprotein component 16/major fat-globule membrane protein/MFG-E8 homolog - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 07-Feb-1997
 C:Accession: E48394
 R:Mather, I.H.; Banoghart, L.R.; Lane, W.S.
 Biochem. Mol. Biol. Int. 29, 545-554, 1993
 A:Title: The major fat-globule membrane proteins, bovine components 15/16 and guinea-
 it-like sequences.
 A:Reference number: A48394; MUID:93250576
 A:Accession: E48394
 A>Status: Preliminary
 A:Molecule type: Protein
 A:Residues: 1-7 <MAT>
 A:Experimental source: milk
 A:Note: sequence extracted from NCBI backbone (NCBIP:131450)
 C:Keywords: glycoprotein

Query Match 32.6%; Score 15; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WA 8
1 1
DB 1 WA 2

RESULT 9
JH0253

gut pentapeptide - Japanese eel
C:Species: Anguilla japonica (Japanese eel)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 11-Apr-1995
C:Accession: JH0253
R:Uesaka, T.; Ikeda, T.; Kubota, T.; Muneoka, Y.; Ando, M.
Biochem. Biophys. Res. Commun. 180, 828-832, 1991
A:Title: Structure and function of a pentapeptide isolated from the gut of the eel.
A:Reference number: JH0253; MUID:92062113
A:Accession: JH0253
A:Molecule type: protein
A:Residues: 1-5 <DB>
A:Experimental source: gut
C:Comment: This peptide increased basal tone of the circular muscle of the esophagogastric, and of the circular muscle of the gastro-intestinal junction.

Query Match 30.4%; Score 14; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 YW 7
1 1
DB 2 FW 3

RESULT 10
B35640

cerebellar degeneration-related protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 24-Jun-1993
C:Accession: B35640
R:Chen, Y.T.; Rettig, W.J.; Yenmandra, A.K.; Kozak, C.A.; Chagantl, R.S.K.; Posner, J.H.
Proc. Natl. Acad. Sci. U.S.A. 87, 3077-3081, 1990
A:Title: Cerebellar degeneration-related antigen: a highly conserved neuroectodermal marker
A:Reference number: A35640; MUID:90222173
A:Accession: B35640
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-6 <CH>

Query Match 30.4%; Score 14; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 YW 7
1 1
DB 1 FW 2

RESULT 11
S21230

dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
C:Species: Phyllomedusa bicolor (two-colored leaf frog)
C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
C:Accession: S21230
R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.
FEBS Lett. 302, 151-154, 1992
A:Title: Identification and characterization of two dermorphins from skin extracts of the
A:Reference number: S21152; MUID:92339502
A:Accession: S21230
A:Status: preliminary
A:Molecule type: protein

A:Residues: 1-7 <MG>
C:Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

Query Match 30.4%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 YW 7
1 1
DB 3 FW 4

RESULT 12
S33567

tubulin beta-3 chain - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
C:Accession: S33567
R:Hinz, U.; Wolk, A.; Renkawitz-Pohl, R.
Development 116, 543-554, 1992
A:Title: Ultrabithorax is a regulator of beta-3 tubulin expression in the Drosophila
A:Reference number: S33567; MUID:93170162
A:Accession: S33567
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-7 <HN>
A:Cross-references: EMBL:X68393; NID:97669; PIDN:CAA8459.1; PID:94377448
C:Genetics:
A:Gene: FlyBase:fbg
A:Cross-references: FlyBase:FBgn0003888
A:introns: 5/3

Query Match 30.4%; Score 14; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 YW 7
1 1
DB 6 FW 7

RESULT 13
S22428

chitin-binding protein - potato (fragment)
C:Species: Solanum tuberosum (potato)
C:Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C:Accession: S22428
R:Millar, D.J.; Allen, A.K.; Smith, C.G.; Slobottom, C.; Slobas, A.R.; Bolwell, G.P.
Biochem. J. 283, 813-821, 1992
A:Title: Chitin-binding proteins in potato (Solanum tuberosum L.) tuber. Characteriza
A:Reference number: S21288; MUID:92272683
A:Accession: S22428
A:Molecule type: protein
A:Residues: 1-8 <ML>
C:Function:
A:Description: may be involved in plant defence
C:Keywords: glycoprotein; hydroxyproline

Query Match 30.4%; Score 14; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 PKLUTY 6
1 1
DB 1 PKDMVY 6

RESULT 14
I49424

cytotoxic T-lymphocyte proteinase 3 (EC 3.4.21.-) - western wild mouse (fragment)

C:Species: Mus spretus (western wild mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I49424
R:Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maizaki, Y.; Nadeau, J.H.
Mamm. Genome 5, 349-355, 1994
A:Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
A:Reference number: I48934; MUID:94319082
A:Accession: I49424
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6 <RES>
A:Cross-references: EMBL:U05745; NID:9497084; PIDD: AAB60481.1; PID:9642831
C:Keywords: hydrolase; serine proteinase

Query Match 28.3%; Score 13; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KLL 4
|||
Db 4 KLL 6

RESULT 15
S33246
neuromodulatory peptide Wkamide-3 - giant African snail
C:Species: Achatina fulica (giant African snail)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Jul-1997
C:Accession: S33246
R:Minakata, H.; Ikeda, T.; Muneoka, Y.; Kobayashi, M.; Nomoto, K.
FEBS Lett. 323, 104-108, 1993
A:Title: Wkamide-1, -2 and -3: novel neuromodulatory peptides isolated from ganglia of t
A:Reference number: S33244; MUID:93265912
A:Accession: S33246
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <MIN>

Query Match 28.3%; Score 13; DB 2; Length 7;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 KLLYW 7
|:|
Db 2 KQMSW 7

Search completed: July 15, 2002, 13:39:06
Job time: 236 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:39:10 ; Search time 10.16 Seconds
(without alignments)
30.488 Million cell updates/sec

Title: US-09-712-819A-4
Sequence: 46
1 PKLIYMA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 148

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	37.0	7	1	GFRP_MOUSE
2	15	32.6	5	1	BPP7_BOTIN
3	13	28.3	7	1	WMA2_ACHFU
4	13	28.3	8	1	ACI_THUAI
5	13	28.3	8	1	LMT2_LOCM1
6	13	28.3	8	1	LPR_LEUMA
7	13	28.3	8	1	PKR2_PBRAM
8	13	28.3	8	1	PKR3_PBRAM
9	12	26.1	7	1	TVS1_LITRU
10	12	26.1	7	1	UP04_MOUSE
11	12	26.1	7	1	WMA3_ACHFU
12	12	26.1	8	1	RS7_MYCIT
13	11	23.9	4	1	OCF3_OCTMI
14	11	23.9	5	1	PSK_DANCA
15	11	23.9	5	1	UF01_MOUSE
16	11	23.9	6	1	EI01_LITRU
17	11	23.9	6	1	L0K1_LOCM1
18	11	23.9	7	1	EI05_LITRU
19	11	23.9	7	1	HY7_PIG
20	11	23.9	7	1	WMA1_ACHFU
21	11	23.9	8	1	AKHG_GRYAT
22	11	23.9	8	1	AKH_LITBAU
23	11	23.9	8	1	AKH_MEIML
24	11	23.9	8	1	AKH_TABAT
25	11	23.9	8	1	CCRN_MACEU
26	11	23.9	8	1	FAR1_PANRE
27	11	23.9	8	1	FAR2_ASCSU
28	11	23.9	8	1	HTF1_PBRAM
29	11	23.9	8	1	HTF2_PBRAM
30	11	23.9	8	1	HTF_TEMO
31	11	23.9	8	1	LCK1_LEUMA
32	11	23.9	8	1	LCK2_LEUMA
33	11	23.9	8	1	LCK3_LEUMA

34	11	23.9	8	1	LCK4_LEUMA	P21143 leucophaea
35	11	23.9	8	1	LCK5_LEUMA	P19987 leucophaea
36	11	23.9	8	1	LCK6_LEUMA	P19988 leucophaea
37	11	23.9	8	1	LCK7_LEUMA	P19989 leucophaea
38	11	23.9	8	1	LCK8_LEUMA	P19990 leucophaea
39	11	23.9	8	1	PLP_BRANA	P81707 brassica na
40	11	23.9	8	1	RPCH_PANBO	P08939 pandanus bo
41	10	21.7	7	1	CARP_MYTEP	P10420 mytilus edu
42	10	21.7	7	1	CCF1_ENTFA	P20104 enterococcu
43	10	21.7	8	1	AL16_CYDPO	P82157 cydia pomon
44	10	21.7	8	1	ANG2_BOTJA	Q10582 bothrops ja
45	10	21.7	8	1	UPAA_HUMAN	P30096 homo sapien

ALIGNMENTS

RESULT 1	GFRP_MOUSE	STANDARD;	PRT;	7 AA.
ID	GFRP_MOUSE			
AC	P99025:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	GTP cyclohydrolyase I feedback regulatory protein (P35) (Fragment).			
GN	GCHFR OR GFRP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Liver;			
RA	Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Van J.X.,			
RA	Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,			
RA	Cowhorne M.;			
RL	Submitted (Aug-1998) to the SWISS-PROT data bank.			
CC	-!- FUNCTION: MEDIATES TERTRAHYDROBIOPTERIN INHIBITION OF GTP			
CC	CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLANILINE			
CC	(BY SIMILARITY).			
CC	-!- SUBUNIT: HOMODIMER (BY SIMILARITY).			
CC	SWISS-2DPAGE; P99025; MOUSE.			
FT	INIT_MET 0			
FT	NON_TER 7			
SO	SEQUENCE 7 AA: 806 MW: 71B5B057273B4700 CRC64;			
Query Match	37.0%; Score 17; DB 1; Length 7;			
Best Local Similarity	80.0%; Pred. NO. 1e+05;			
Matches	4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1 PKLI 5			
DB	1 PILLI 5			
RESULT 2	BPP7_BOTIN	STANDARD;	PRT;	5 AA.
ID	BPP7_BOTIN			
AC	P30425:			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-FEB-1994 (Rel. 28, Last annotation update)			
DE	Bradykinin-potentiating peptide S5.2 (5A) (Angiotensin-converting			
DE	enzyme inhibitor).			
OS	Bothrops insularis (Island jararaca) (Queimada jararaca).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Leptidosauria; Squamata; ScleroGLOSSA; Serpentes; Colubroidae;			
OC	Viperidae; Crotalinae; Bothrops.			
OX	NCBI_TaxID=8723;			
RP	SEQUENCE.			
RC	TISSUE=Venom;			

RX MEDLINE=90351557; PubMed=2386615;
 RA Caltura A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating
 RT peptides from Bothrops insularis snake venom."
 RL J. Protein Chem. 9:221-227(1990).
 CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
 DR PIR: G37196; G37196.
 KW Hypotensive agent; Venom.
 FT MOD_RES 1
 SQ SEQUENCE 5 AA; 629 MW; 776DC37326B0000 CRC64;

Query Match 32.6%; Score 15; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 WA 8
 DB 3 WA 4

RESULT 3
 ID WMA2_ACHFV STANDARD; PRT; 7 AA.
 AC P35920;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Wmamide-2.
 OS Achatina fulica (Giant African snail).
 CC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 CC Achatinacea; Achatinidae; Achatina.
 CX NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Ganglion;
 RX MEDLINE=93265912; PubMed=8495720;
 RA Mlaketa H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
 RT "Wmamide-1, -2 and -3: novel neuromodulatory peptides isolated from
 RT ganglia of the African giant snail, Achatina fulica."
 RL FEBS Lett. 323:104-108(1993).
 KW Neuropeptide; Amidation.
 FT MOD_RES 7
 SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 28.3%; Score 13; DB 1; Length 7;
 Best Local Similarity 33.3%; Pred. No. 1e+05;
 Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 KLIYV 7
 DB 2 KOMSVW 7

RESULT 4
 ID ACI_THUOL STANDARD; PRT; 8 AA.
 AC P18691;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-NOV-1990 (Rel. 16, Last annotation update)
 DE Angiotensin-converting enzyme inhibitor.
 OS Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Acanthomorpha; Acanthopterygii; Perciformes; Scombroidei;
 CC Scombridae; Thunnus.
 CX NCBI_TaxID=8236;
 RN [1]

RP SEQUENCE.
 RC TISSUE=Muscle;
 RX MEDLINE=88326322; PubMed=2415688;
 RA Kohma Y., Matsumoto S., Oka H., Teramoto T., Okabe M., Mimura T.;
 RT "Isolation of angiotensin-converting enzyme inhibitor from tuna
 RT muscle."
 RL Biochem. Biophys. Res. Commun. 155:332-337(1988).
 DR PIR: A31570; A31570.
 SQ SEQUENCE 8 AA; 953 MW; 6AA863733051F1B7 CRC64;

Query Match 28.3%; Score 13; DB 1; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 IYW 7
 DB 4 IKW 6

RESULT 5
 ID LMT2_LOCM1 STANDARD; PRT; 8 AA.
 AC P22396;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-AUG-1991 (Rel. 19, Last annotation update)
 DE Locustamyotropin 2 (LOW-MT-2).
 OS Locusta migratoria (Migratory locust).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 CC Acridomorpha; Acridoidea; Acrididae; Locusta.
 CX NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RA Schoofs L., Holman G.W., Hayes T.K., Nachman R.J., de Loof A.;
 RT "Isolation, identification and synthesis of locustamyotropin II, an
 RT additional neuropeptide of locusta migratoria. Member of the
 RT cephalomyotropic peptide family."
 RL Insect Biochem. 20:479-484(1990).
 CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
 CC (MYOTROPIC ACTIVITY).
 CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
 DR InterPro: IPR001484; Pyrokinin.
 DR PROSITE: PS00539; PYROKININ; 1.
 KW Neuropeptide; Amidation; Pyrokinin.
 FT MOD_RES 8
 SQ SEQUENCE 8 AA; 934 MW; 26341771A9CA87B CRC64;

Query Match 28.3%; Score 13; DB 1; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKL 3
 DB 6 PRL 8

RESULT 6
 ID LPK_LEUMA STANDARD; PRT; 8 AA.
 AC P13049;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Leukopyrokinin (LPK) (LEU-PK).
 OS Leucophaea maderae (Madeira cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Orthopteroidea; Dicopteroidea; Blattaria;
 CC Blaberidae; Blaberidae; Leucophaea.
 CX NCBI_TaxID=6988;

```

RN [1]
RP SEQUENCE.
RX MEDLINE=86269041; PubMed=3015140;
RA Nachman R.J., Holman G.M., Cook B.J.;
RT "Active fragments and analogs of the insect neuropeptide
RT leucopyrokinin: structure-function studies.,"
RL Biochem. Biophys. Res. Commun. 137:936-942(1986).
RN [2]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE-Head:
RX MEDLINE=87052651; PubMed=2877794;
RA Holman G.M., Cook B.J., Nachman R.J.;
RT "Primary structure and synthesis of a blocked myotropic
RT neuropeptide isolated from the cockroach, Leucophaea maderae.,"
RL Comp. Biochem. Physiol. 85C:219-224(1986).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -1- MISCELLANEOUS: AN ANALOG WITHOUT THE N-TERMINAL PCA RESIDUE WAS
CC SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (144%) THAN THE
CC PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE OF LPK MOST
CC CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED TO THE
CC PENTAPEPTIDE FRAGMENT TPRL.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro: IPR001484; PYROKININ.
DR InterPro: IPR001484; PYROKININ.
DR PROSITE: PS00539; PYROKININ.1.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA: 949 MW: 92341771AD5A1B6 CRC64:

```

```

Query Match 28.3%; Score 13; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PKL 3
   1:1
DB 6 PRL 8

```

```

RESULT 7
PRL3_PERAM ID STANDARD; PRT; 8 AA.
AC P82692;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pyrokinin-2 (Pep-PK-2) (FXPRL-amide).
OS Periplaneta americana (American cockroach).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
CC Blattellidae; Blattidae; Periplaneta.
OK NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-Corpora cardiaca;
RX MEDLINE=97353923; PubMed=9210163;
RA Predel R., Kellner R., Kaufman R., Penzlin H., Gaede G.;
RT "Isolation and structural elucidation of two pyrokinins from the
RT retrocerebral complex of the American cockroach.,"
RL Peptides 18:473-478(1997).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.,"
RL J. Comp. Neurol. 419:352-363(2000).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -1- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA: 997 MW: 0B34177409D772C7 CRC64:

```

```

CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR InterPro: IPR001484; PYROKININ.
DR PROSITE: PS00539; PYROKININ; FALSE NEG.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA: 884 MW: C834176DD9D77775 CRC64:

```

```

Query Match 28.3%; Score 13; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PKL 3
   1:1
DB 6 PRL 8

```

```

RESULT 8
PRL3_PERAM ID STANDARD; PRT; 8 AA.
AC P82618;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pyrokinin-3 (FXPRL-amide).
OS Periplaneta americana (American cockroach).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
CC Blattellidae; Blattidae; Periplaneta.
OK NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-Retrocerebral complex;
RX MEDLINE=99212469; PubMed=1016736;
RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;
RT "Differential distribution of pyrokinin-isoforms in cerebral and
RT abdominal neurohemal organs of the American cockroach.,"
RL Insect Biochem. Mol. Biol. 29:139-144(1999).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.,"
RL J. Comp. Neurol. 419:352-363(2000).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC (MYOTROPIC ACTIVITY).
CC -1- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -1- MASS SPECTROMETRY: MW=996.5; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA: 997 MW: 0B34177409D772C7 CRC64:

```

```

Query Match 28.3%; Score 13; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PKL 3
   1:1
DB 6 PRL 8

```

```

RESULT 9
TY51_LITRU ID STANDARD; PRT; 7 AA.
AC P82065;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Typtophyllin 5.1.
OS Litorea rubella (Desert tree frog).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OC NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;
 RA Steinboerner S.T., Wabnitz P.A., Maugh R.J., Bowle J.H., Gao C.,
 RT Tyler M.J., Wallace J.C.;
 RT 'The structure of new peptides from the Australian red tree frog
 RT 'Litoria rubella'. The skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians.';
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: MAY ACT AS A NEUROMODULATOR OR NEUOTRASMITTER.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=965; METHOD=FM.
 KW Amphibian skin; Amidation; Neuropeptide.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 7 7 AMIDATION.
 SO SEQUENCE 7 AA; 983 MW; 7401E9D3676046B0 CRC64;

Query Match 26.1%; Score 12; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ITW 7
 1 1
 DB 2 IPW 4

RESULT 10
 UFG4_MOUSE
 ID UFG4_MOUSE STANDARD; PRT; 7 AA.
 AC P38642;
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P46) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.0, ITS MW IS: 46 kDa.
 FT NON_TER 7 7
 SO SEQUENCE 7 AA; 766 MW; 68640AB77632700 CRC64;

Query Match 26.1%; Score 12; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PK 2
 1 1
 DB 1 PK 2

RESULT 11
 WMA3_ACHFU
 ID WMA3_ACHFU STANDARD; PRT; 7 AA.
 AC P35921;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Wamide-3.
 OS Achatina fulica (Giant African snail).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
 OC Achatinacea; Achatinidae; Achatina.
 OC NCBI_TaxID=6530;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Ganglion;
 RX MEDLINE=93265912; PubMed=8495720;
 RA Minakata H., Ikeda T., Munneke Y., Kobayashi M., Nomoto K.;
 RT "Wamide-1, -2 and -3: novel neuromodulatory peptides isolated from
 RT ganglia of the African giant snail, Achatina fulica.";
 RL FEBS Lett. 323:104-108(1993).
 DR PIR: S33244; S33244.
 KW Neuropeptide; Amidation.
 FT MOD_RES 7 7 AMIDATION.
 SO SEQUENCE 7 AA; 965 MW; 7362D5B69B132310 CRC64;

Query Match 26.1%; Score 12; DB 1; Length 7;
 Best Local Similarity 33.3%; Pred. No. 1e+05;
 Matches 2; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 KLIW 7
 1 1
 DB 2 KEMSW 7

RESULT 12
 RS7_MYCIT
 ID RS7_MYCIT STANDARD; PRT; 8 AA.
 AC P33564;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE 30S ribosomal protein S7 (Fragment).
 GN RPSG.
 OS Mycobacterium intracellulare.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93197130; PubMed=8451173;
 RA Nair J., Rouse D.A., Morris S.L.;
 RT "Nucleotide sequence analysis of the ribosomal S12 gene of
 RT Mycobacterium intracellulare.";
 RL Nucleic Acids Res. 21:1039-1039(1993).
 CC -1- FUNCTION: PROTEIN S7 BINDS SPECIFICALLY TO PART OF THE 3' END OF
 CC 16S RIBOSOMAL RNA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE S7P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: L08171; AAA25376.1; -.
 CC DR PIR: S3538; S3538.
 CC DR InterPro: IPR000235; Ribosomal_S7.
 CC DR PROSITE: PS00052; RIBOSOMAL_S7; PARTIAL.
 KW Ribosomal protein; rRNA-binding.
 FT INT_MET 8 8 BY SIMILARITY.
 FT NON_TER 8 8
 SO SEQUENCE 8 AA; 850 MW; 63276DC768732417 CRC64;

Query Match 26.1%; Score 12; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PK 2
11
Db 7 PK 8

RESULT 13

OCPI3_OCPMI STANDARD; PRT; 4 AA.

AC P58649;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cardiac peptide Ocp-3/Ocp-4.
OS Octopus minor (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=89766;

RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.

RC TISSUE=Brain;

RX PubMed=10876044;

RA Iwakoshi E., Hisada M., Minakata H.;

RT "Cardiac peptides isolated from the brain of a Japanese octopus,
Octopus minor.";

RL Peptides 21:623-630(2000).

DE FUNCTION: Cardiac peptide; has both positive chronotropic and
inotropic effects on the heart. Ocp-4 is a 1000 time less

active than Ocp-3.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- PTM: Ocp-4 has D-Ser instead of L-Ser.

CC -1- MASS SPECTROMETRY: MW=395.2; METHOD=MALDI.

KV Hormone: D-amino acid.

FT MOD_RES 2 D-SERINE (IN OCP-4);

SQ SEQUENCE 4 AA; 463 MW; 6A8365B810000000 CRC64;

Query Match 23.9%; Score 11; DB 1; Length 4;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 W 7
1
Db 3 W 3

RESULT 14

PSK_DAUCA STANDARD; PRT; 5 AA.

AC P58261;
DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-beta)].

OS Daucus carota (Carrot).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliopsida; core eudicots;

OC Asteridae; eusterids II; Apiales; Apiaceae; Daucus.

OX NCBI_TaxID=4039;

RP SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.

RC STRAIN=cv. US-Harumakigosun;

RA MEDLINE=20212743; PubMed=10750705;

RA Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
Kamada H., Sakagami Y.;

RT "A secreted growth factor, phytosulfokine, acting as a
stimulatory factor of carrot somatic embryo formation.";

RL Plant Cell Physiol. 41:27-32(2000).

DE FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
EMBRYOS.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
CC PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
KW Growth factor; Sulfation.
FT PEPTIDE 1 4 PHYTOSULFOKINE-BETA.
FT MOD_RES 1 1 SULFATION.
FT MOD_RES 3 3 SULFATION.
SQ SEQUENCE 5 AA; 687 MW; 76C1BB504B300000 CRC64;

Query Match 23.9%; Score 11; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 IY 6
11
Db 2 IY 3

RESULT 15

UF01_MOUSE STANDARD; PRT; 5 AA.

AC P38639;
DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-FEB-1995 (Rel. 31, Last annotation update)

DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RP SEQUENCE.

RC TISSUE=Fibroblast;

RX MEDLINE=95009907; PubMed=7523108;

RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;

RT "Separation and sequencing of familial and novel murine proteins
using preparative two-dimensional gel electrophoresis.";

RL Electrophoresis 15:735-745(1994)

CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
PROTEIN IS: 6.6, ITS MW IS: 19 KDa.

CC NON_TER 5 5

SQ SEQUENCE 5 AA; 717 MW; 7364087043100000 CRC64;

Query Match 23.9%; Score 11; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 W 7
1
Db 1 W 1

Search completed: July 15, 2002, 13:42:21
Job time: 191 sec

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHEICAL 1.0 KDA PROTEIN (FRAGMENT).
OS Coxiella burnetii.
OC Bacteria: Proteobacteria: gamma subdivision: Legionellaceae group:
OC Coxiella group: Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE PHASE I;
RX MEDLINE=98348442; PubMed=9683477;
RA Williams H., Jaeger C., Baljer G.;
RT *Physical and genetic map of the obligate intracellular bacterium
RT Coxiella burnetii.*;
RL J. Bacteriol.180:3816-3822(1998).
DR EMBL: AF064963; AAD09947.1; -.
KM Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 8 AA: 993 MW: 04655A453772727 CRC64;

Query Match 34.8%; Score 16; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIYW 7
: | |
DB 1 LLLW 4

RESULT 3
ID 094VJ4 PRELIMINARY; PRT; 8 AA.
AC 094VJ4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Varanus bengalensis nebulosis.
OC Mitochondrion.
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Lepidosauria: Squamata: Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169827;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RT Cladistics 17:0-0(2001).
DR EMBL: AF407492; AAL10031.1; -.
DR EMBL: AF407492; AAL10031.1; -.
KM Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA: 1053 MW: E8B5B9C733640056 CRC64;

Query Match 32.6%; Score 15; DB 8; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIYW 7
: | |
DB 1 MIRM 4

RESULT 4
ID 094V82 PRELIMINARY; PRT; 8 AA.
AC 094V82;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
GN COI.
OS Varanus yuwonoi.

OG Mitochondrion.
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Lepidosauria: Squamata: Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX NCBI_TaxID=169856;
RN [1]
RP SEQUENCE FROM N.A.
RA Ast J.C.;
RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RT Cladistics 17:0-0(2001).
DR EMBL: AF407535; AAL10157.1; -.
KM Mitochondrion.
FT NON_TER 8
SQ SEQUENCE 8 AA: 1045 MW: EFC775A6C3640056 CRC64;

Query Match 32.6%; Score 15; DB 8; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIYW 7
: | |
DB 1 MIRM 4

RESULT 5
ID 042507 PRELIMINARY; PRT; 8 AA.
AC 042507;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)
DE HEAT SHOCK PROTEIN (FRAGMENT).
GN HSP70C OR HSP70A OR HSP70B.
OC Triticum aestivum (wheat).
OC Eukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: Liliopsida; Poales; Poaceae; Poaceae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV MUSTANG;
RX MEDLINE=96189275; PubMed=8605312;
RA Joshi C.P., Kumar S., Nguyen H.T.;
RT "Application of modified differential display technique for cloning
RT and sequencing of the 3' region from three putative members of wheat
RT HSP70 gene family."; Plant Mol. Biol. 30:641-646(1996).
RL Plant Mol. Biol. 30:641-646(1996).
DR EMBL: L41507; AAB02333.1; -.
DR EMBL: L41505; AAB02331.1; -.
DR EMBL: L41506; AAB02332.1; -.
KM Heat shock.
FT NON_TER 1
SQ SEQUENCE 8 AA: 886 MW: 71B2CB1B10532768 CRC64;

Query Match 30.4%; Score 14; DB 10; Length 8;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKL 3
: | |
DB 2 PKL 4

RESULT 6
ID 050556 PRELIMINARY; PRT; 7 AA.
AC 050556;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE GLYA (FRAGMENT).
GN GLYA.

OS Actinobacillus actinomycetemcomitans (Haemophilus
OC actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33384;
RX MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans.";
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -.
FT NON_TER
SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 28.3%; Score 13; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KLLY 6
DB 2 RLPVY 6

RESULT 7
ID 066205 PRELIMINARY; PRT; 7 AA.
AC 066205;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE MEMBRANE PROTEIN (1 IS 3RD BASE IN CODON) (FRAGMENT).
OS Transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FS772/70;
RX MEDLINE=88216185; PubMed=2835592;
RA Britton P., Carmenes R.S., Page K.W., Garves D.J., Parra F.;
RT "Sequence of the Nucleoprotein Gene from a Virulent British Field
RT Isolate of Transmissible Gastroenteritis Virus and its Expression in
RT Saccharomyces Cerevisiae.";
RL Mol. Microbiol. 2:89-99(1988).
DR EMBL; Y00542; CAA68606.1; -.
FT NON_TER
SQ SEQUENCE 7 AA; 869 MW; 72C691E727233B70 CRC64;

Query Match 28.3%; Score 13; DB 12; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLL 4
DB 2 KLL 4

RESULT 8
ID 087471 PRELIMINARY; PRT; 8 AA.
AC 087471;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE HIFA (FRAGMENT).
GN HIFA.

OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ERGAN (EIA);
RX MEDLINE=96389689; PubMed=9721313;
RA Mhlanga-Mutlangadura T., Morlin G., Smith A.L., Eisenstark A.,
RA Golomb M.;
RT "Evolution of the major plus gene cluster of haemophilus
RT influenzae.";
RT J. Bacteriol. 180:4693-4703(1998).
RL EMBL; AF071762; AAC35830.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 876 MW; DAB44451A7272325 CRC64;

Query Match 28.3%; Score 13; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KLL 4
DB 2 KLL 4

RESULT 9
ID 002831 PRELIMINARY; PRT; 8 AA.
AC 002831;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PRO ALPHA 1 TYPE III COLLAGEN PROTEIN (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96377339; PubMed=8783186;
RA Metaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA Vuorio E.;
RT "Evidence for insufficient chondrocytic differentiation during repair
RT of full-thickness defects of articular cartilage.";
RL Matrix Biol. 15:39-47(1996).
DR EMBL; S83571; AAD14433.1; -.
FT NON_TER
SQ SEQUENCE 8 AA; 1028 MW; B859C7272EA77371 CRC64;

Query Match 28.3%; Score 13; DB 6; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 YW 7
DB 1 HW 2

RESULT 10
ID P92421 PRELIMINARY; PRT; 7 AA.
AC P92421; P92419;
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-JAN-1998 (TREMblrel. 05, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS psathyrostachys fragilis.
OG Chloroplast.

OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Trilicaceae; Psathyrostachys.
 OX NCBI_TaxID=37729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H4372, AND H917; TISSUE=LEAVES;
 RA Petersen G., Seberg O.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 277753; CAB01337.1; -
 DR EMBL: 277752; CAB01334.1; -
 KW Ribosomal protein; Chloroplast.
 FT NON_TER 1
 SQ SEQUENCE 7 AA: 894 MW: 673414033277700 CRC64;

Query Match 26.1%; Score 12; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PK 2
 II
 DB 2 PK 3

RESULT 11
 ID P92385 PRELIMINARY; PRT: 7 AA.
 AC P92385; P92383;
 DT 01-MAY-1997 (TREMUREL. 03, Created)
 DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
 DT 01-JAN-1998 (TREMUREL. 05, Last annotation update)
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
 GN RPS11.
 OS Hordeum maritimum (Seaside barley).
 OC Chloroplast.
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Trilicaceae; Hordeum.
 OX NCBI_TaxID=4519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H299, AND H801; TISSUE=LEAVES;
 RA Petersen G., Seberg O.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 277763; CAB01367.1; -
 DR EMBL: 277762; CAB01364.1; -
 KW Ribosomal protein; Chloroplast.
 FT NON_TER 1
 SQ SEQUENCE 7 AA: 894 MW: 673414033277700 CRC64;

Query Match 26.1%; Score 12; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PK 2
 II
 DB 2 PK 3

RESULT 12
 ID P92372 PRELIMINARY; PRT: 7 AA.
 AC P92372;
 DT 01-MAY-1997 (TREMUREL. 03, Created)
 DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
 GN RPS11.
 OS Haynaldia villosa.
 OC Chloroplast.
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Haynaldia.
 OX NCBI_TaxID=40247;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H5561; TISSUE=LEAF;
 RX MEDLINE=97271648; PubMed=9126564;
 RA Petersen G., Seberg O.;
 RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
 RT sequence data."
 RL Mol. Phylogenet. Evol. 7:217-230(1997).
 DR EMBL: 277741; CAB01301.1; -
 KW Chloroplast.
 FT NON_TER 1
 SQ SEQUENCE 7 AA: 894 MW: 673414033277700 CRC64;

Query Match 26.1%; Score 12; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PK 2
 II
 DB 2 PK 3

RESULT 13
 ID P92403 PRELIMINARY; PRT: 7 AA.
 AC P92403;
 DT 01-MAY-1997 (TREMUREL. 03, Created)
 DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
 GN RPS11.
 OS Lophopyrum elongatum (Tall wheatgrass) (Argropyrum elongatum).
 OC Chloroplast.
 OC Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Trilicaceae; Lophopyrum.
 OX NCBI_TaxID=4588;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H6692; TISSUE=LEAF;
 RX MEDLINE=97271648; PubMed=9126564;
 RA Petersen G., Seberg O.;
 RT "Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA
 RT sequence data."
 RL Mol. Phylogenet. Evol. 7:217-230(1997).
 DR EMBL: 277743; CAB01307.1; -
 KW Chloroplast.
 FT NON_TER 1
 SQ SEQUENCE 7 AA: 894 MW: 673414033277700 CRC64;

Query Match 26.1%; Score 12; DB 8; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PK 2
 II
 DB 2 PK 3

RESULT 14
 ID P92425 PRELIMINARY; PRT: 7 AA.
 AC P92425;
 DT 01-MAY-1997 (TREMUREL. 03, Created)
 DT 01-MAY-1997 (TREMUREL. 03, Last sequence update)
 DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
 DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
 GN RPS11.

```

OS Pseudoroegneria spicata.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Pseudoroegneria.
OX NCBI_TaxID=4604;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H9082; TISSUE-LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G.; Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77744; CAB01310.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

```

```

Query Match          26.1%; Score 12; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PK 2
   ||
DB 2 PK 3

```

```

RESULT 15
P92387          PRELIMINARY;      PRT;      7 AA.
AC P92387:
DT 01-MAY-1997 (TRENBLER. 03, Created)
DT 01-MAY-1997 (TRENBLER. 03, Last sequence update)
DT 01-DEC-2001 (TRENBLER. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Henrardia persica.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Henrardia.
OX NCBI_TaxID=37678;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H9556; TISSUE-LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G.; Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77748; CAB01322.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

```

```

Query Match          26.1%; Score 12; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PK 2
   ||
DB 2 PK 3

```

Search completed: July 15, 2002, 13:42:04
Job time: 194 sec

10

11

12

13

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:27:15 ; Search time 28.83 Seconds
(without alignments)
30.822 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46

Sequence: 1 PKLITVWA 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 69368

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Database :

Listing first 45 summaries

A_Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	91.3	8	20	AA140668
2	42	91.3	8	21	AA140668
3	31	67.4	8	20	AA140669
4	31	67.4	8	21	AA140669
5	28	60.9	8	20	AA140667
6	28	60.9	8	21	AA140667
7	26	56.5	8	19	AA137187
8	26	56.5	8	19	AA137187
9	25	54.3	6	22	AA182651
10	25	54.3	6	22	AA182651
11	25	54.3	7	22	AA182651

12	25	54.3	7	22	AA182651
13	24	52.2	4	22	AA182651
14	24	52.2	6	22	AA182652
15	24	52.2	8	14	AA182652
16	23	50.0	7	22	AA182652
17	22	47.8	5	4	AA182652
18	22	47.8	5	14	AA182652
19	22	47.8	5	14	AA182652
20	22	47.8	5	17	AA182652
21	22	47.8	5	19	AA182652
22	22	47.8	5	19	AA182652
23	22	47.8	5	19	AA182652
24	22	47.8	5	19	AA182652
25	22	47.8	5	19	AA182652
26	22	47.8	6	4	AA182652
27	22	47.8	6	4	AA182652
28	22	47.8	6	4	AA182652
29	22	47.8	6	4	AA182652
30	22	47.8	6	16	AA182652
31	22	47.8	6	21	AA182652
32	22	47.8	7	22	AA182652
33	22	47.8	7	22	AA182652
34	22	47.8	8	18	AA182652
35	22	47.8	8	19	AA182652
36	22	47.8	8	19	AA182652
37	22	47.8	8	20	AA182652
38	22	47.8	8	21	AA182652
39	21	45.7	6	17	AA182652
40	21	45.7	6	22	AA182652
41	21	45.7	8	11	AA182652
42	21	45.7	8	11	AA182652
43	21	45.7	8	18	AA182652
44	21	45.7	8	18	AA182652
45	21	45.7	8	20	AA182652

ALIGNMENTS

RESULT 1	
AA140668	AA140668 standard; peptide; 8 AA.
XX	
AC	AA140668;
XX	
DT	01-DEC-1999 (first entry)
XX	
DE	A2 derivative #10, beta strand of scaffold protein structure.
XX	
KW	Scaffold protein; beta strand; stabilize antigen; vaccine; tumour; chemotherapeutic agent.
XX	
OS	Synthetic.
XX	
PN	EP947582-A1.
XX	
PD	06-OCT-1999.
XX	
PF	31-MAR-1998; 98EP-0870065.
XX	
PR	31-MAR-1998; 98EP-0870065.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX	
DR	WPI; 1999-542958/46.
XX	
PT	New scaffold protein, useful for stabilizing antigens used as vaccines
XX	
PS	Disclosure; Page 5; 105pp; English.
XX	

Collagenase cleave
Human SNP related
All-D peptide used
Sample peptide to
H11 binding site c
Peptide having pit
GHRP-6 (Growth hor
GHRP-6 analog. sy
Growth hormone rel
Peptide determined
Peptide determined
Growth hormone rel
K. oxytoca R-spect
Peptide having pit
Peptide having pit
Peptide having pit
Peptide having pit
Cyclo(-Tyr-Lrp-Ala
Human HBRG protein
H11 binding site c
H11 binding site c
Protein kinase C r
Human presentin I
Cyclomaltodextrin
T cell epitope/MHC
Somatostatin pepti
Peptide motif need
Rhesus D antibody
Accessory moiety d
TMD5 region amino
Anti-fungal peptid
Protein kinase C r
Antigenic site of

CC Sequences AAY40659-Y40679 are functionally equivalent derivatives of the
CC A2 peptide (AAY40604) which forms part of a scaffold protein. A2 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

SO Sequence 8 AA:

Query Match 91.3%; Score 42; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 7
| | | | | | |
DB 2 PKLLIYW 8

RESULT 2

AAB30007 AAB30007 standard; Peptide: 8 AA.

AC AAB30007;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S2 peptide SEQ ID NO: 68.

XX Human: CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN MO20060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99MO-EP02283.

PR 01-APR-1999; 99MO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments -

PS Disclosure; Page 14; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.

SO Sequence 8 AA:

Query Match 91.3%; Score 42; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 7
| | | | | | |
DB 2 PKLLIYW 8

RESULT 3

AAY40669 AAY40669 standard; peptide: 8 AA.

AC AAY40669;

DT 01-DEC-1999 (first entry)

DE A2 derivative #11, beta strand of scaffold protein structure.

XX Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.

OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines
PT -
PS Disclosure; Page 5; 105pp; English.

CC Sequences AAY40659-Y40679 are functionally equivalent derivatives of the
CC A2 peptide (AAY40604) which forms part of a scaffold protein. A2 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will

CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX Sequence 8 AA;

Query Match 67.4%; Score 31; DB 20; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIY 6
Db 2 PKLLIY 7

RESULT 4

AAB30008 AAB30008 standard; Peptide: 8 AA.

AC AAB30008;
DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S2 peptide SEQ ID NO: 69.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO20006070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

XX Scaffold composed of single-chain polypeptide having beta sandwich
XX architecture carrying new and randomized peptide sequences useful as
XX supporting framework and carrying antigen- or receptor binding
XX fragments -
XX
XX Disclosure: Page 14; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
XX based upon the human CTLA-4 SCA domain. These scaffold proteins can be
XX used as a scaffold to bind antigen- or receptor-binding fragments. These
XX can be used in the treatment of diseases such as cancer,
XX atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
XX diabetic retinopathy. Sequences AAB29930-B29939 were used in the
XX production of the proteins of the invention.

SQ Sequence 8 AA;

Query Match 67.4%; Score 31; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIY 6
IIIIII

Db 2 PKLLIY 7

RESULT 5

AAV40667 AAV40667 standard; peptide: 8 AA.

AC AAV40667;

DT 01-DEC-1999 (first entry)

DE A2 derivative #9, beta strand of scaffold protein structure.

KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.

OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PE 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

XX New scaffold protein, useful for stabilizing antigens used as vaccines
XX
XX
XX Disclosure: Page 5; 105pp; English.

XX Sequences AAV40659-Y40679 are functionally equivalent derivatives of the
XX A2 peptide (AAV40604) which forms part of a scaffold protein. A2 is a
XX beta strand peptide which forms part of a beta sheet. Peptides
XX (AAV40601-Y40609) together form a single-chain scaffold protein which
XX contains at least 1 disulfide bond, contains less than 10% alpha helix
XX and contains at least 6 beta-strands. The scaffold protein is constructed
XX of beta strands SI-S6, and may also include beta strands AI-A3, or any
XX functionally equivalent derivative of these sequences. The beta strands
XX form two beta sheets SI/S4/S3 and S6/S5/S2 with each strand connected to
XX the next by hydrogen bonds, which generate a beta sandwich architecture.
XX If the additional beta strands AI-A3 are included in the structure the
XX scaffold is constructed of two beta sheets, with the structures
XX AI/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
XX other via amino acid loops, where at least one of the loops binds to a
XX receptor or antigen. The scaffold protein is used to stabilize antigens
XX or whole proteins such as receptors, or their fragments. It may be used
XX to bind two separate molecules. For example, one surface of the scaffold
XX may be bound to a protein which binds to a tumour antigen. This will
XX target the complex to tumour cells. Another surface may be bound to a
XX cytotoxic molecule or an autoimmune antibody which may then kill the
XX tumour cells. Therefore the scaffold protein may be used to target
XX chemotherapeutic agents to specific cells. It may also be used to
XX stabilize individual peptides in a peptide library and may be used in
XX diagnostic techniques, and to stabilize antigens used as vaccines.

SQ Sequence 8 AA;

Query Match 60.9%; Score 28; DB 20; Length 8;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIY 6
I:IIII
Db 2 PKLLIY 7

```

RESULT 6
AAB30006
ID AAB30006 standard; Peptide: 8 AA.
XX
AC AAB30006;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S2 peptide SEQ ID NO: 67.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200006070-A1.
XX
PD 12-OCT-2000.
XX
PE 01-APR-1999; 99WO-EP02283.
XX
PR 01-APR-1999; 99WO-EP02283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI; 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments -
XX
PS Disclosure: Page 14; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding proteins. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.
XX
SQ Sequence 8 AA;

```

Query Match 60.98; Score 28; DB 21; Length 8;
 Best Local Similarity 83.38; Pred. No. 6.4e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 PKLIIV 6
   1:||||
Db 2 prilly 7

```

```

RESULT 7
AAW37187
ID AAW37187 standard; peptide: 8 AA.
XX
AC AAW37187;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding N-acetylated peptide derivative 6.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.

```

```

XX Key Location/Qualifiers
XX Modified-site 1
XX Modified-site /note= "N-terminal acetyl"
XX Modified-site 8
XX Modified-site /note= "C-terminal amide"
XX
PN WO9801467-A2.
XX
PD 15-JAN-1998.
XX
PE 04-JUL-1997; 97WO-EP03549.
XX
PR 07-APR-1997; 97GB-0007041.
XX
PR 05-JUL-1996; 96GB-0014197.
XX
PA (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX
PA (NOVS) NOVARTIS AG.
XX
PI Boettger A, Boettger V, Chene P, Furet P, Garcia-echeverria C;
PI Hochkeppel H, Lane D, Pickxley S;
XX
DR WPI; 1998-100996/09.
XX
PT Compounds binding to MDM2 protein and inhibit its interaction with
PT p53 - useful in, e.g. diagnosis and treatment of cancer and viral
PT infections and identifying binding agents
XX
PS Example 1; Page 19; 45pp; English.
XX
CC This is a N-acetylated peptide derivative capable of binding to a human
CC oncogenic protein MDM2. The MDM2 binding peptides can specifically
CC inhibit or block the binding of MDM2 to the human p53 protein, in vitro
CC or in vivo. Inhibiting the interaction between the p53 and MDM2 can
CC induce growth arrest or apoptosis in tumour cells comprising a wild-type
CC p53 and non-elevated levels of MDM2. The peptides may be used to identify
CC molecules that bind to MDM2 and to identify and design inhibitors of
CC MDM2/p53 binding. They may also be used to purify binding partners
CC especially MDM2, diagnose disease by measuring levels of MDM2 in blood
CC of cancer and leukemia patients and for treatment or prevention of
CC disease involving p53/MDM2 interactions, especially tumours and viral
CC infections. The peptides can be administered nasally, rectally, orally
CC or by injection. By interfering with MDM2/p53 interaction, the peptides
CC can activate p53 function and accumulation in normal cells. The peptides
CC which mimic the MDM2 binding site in p53, have a significantly greater
CC blocking activity compared with wild-type p53.
XX
SQ Sequence 8 AA;

```

Query Match 56.58; Score 26; DB 19; Length 8;
 Best Local Similarity 42.98; Pred. No. 6.4e+05;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 PKLIIV 7
   1: : ||
Db 1 prfndyw 7

```

```

RESULT 8
AAW37178
ID AAW37178 standard; peptide: 8 AA.
XX
AC AAW37178;
XX
DT 20-JUL-1998 (first entry)
XX
DE Human oncogenic protein MDM2 binding peptide derivative 6.
XX
KW MDM2; oncogenic protein; p53; human; inhibition; interaction; cancer;
KW tumour; diagnosis; binding; viral infection.
XX
OS Synthetic.
OS Homo sapiens.

```


XX 29-NOV-1999; 99US-0168594.
PR 28-NOV-2000; 2000US-0724842.
XX
XX (NEUR-) NEUROCHEM INC.
XX
XX Chalfour R, Hebert L, Kong X, Gervais F;
PI WPI: 2001-441458/47.
XX
XX Preventing/creating amyloid-related disease, especially Alzheimer's
PT disease, comprises administering antigenic all-D peptide, e.g. as
PT vaccine, which elicits production of antibodies to prevent
PT fibrillogenesis and associated cellular toxicity -
XX
XX Disclosure; Page 12; 31pp; English.
XX
XX The present sequence is that of an all-D peptide suitable for
CC use for preparing vaccines for preventing or treating Alzheimer's
CC disease and other amyloid related disorders in humans. It is based
CC on a portion of amyloid-beta peptide (see AAB82622). Vaccines
CC are produced using 'non-self' peptides synthesised from the
CC unnatural D-configuration amino acids to avoid the drawbacks of
CC 'self' proteins. The all-D peptides need not be aggregated to be
CC operative or immunogenic. They preferably interact with at
CC least 1 region of an amyloid protein, e.g. the beta-sheet region
CC or GAG-binding site region, the amyloid-beta peptide, or their
CC immunogenic fragments, protein conjugates, immunogenic derivative
CC peptides and immunogenic peptidomimetics. Examples include all-D
CC peptides corresponding to residues 1-42, 1-40, 1-35, 1-28, 1-7,
CC 10-16, 16-21 and 36-42 of the amyloid-beta peptide and the all-D
CC derivative peptides given in AAB82623-64. The vaccine elicits a
CC preferential TH-2 or TH-1 response, preventing fibrillogenesis and
CC associated cellular toxicity. The amyloid related diseases may be
CC localised amyloidosis, e.g. diabetes type II, neurodegenerative
CC diseases, e.g. bovine spongiform encephalitis, Creutzfeldt-Jakob
CC disease, scrapie, cerebral amyloid angiopathy, and prion protein
CC related disorders, or systemic amyloidosis associated with chronic
CC infection (e.g. tuberculosis) or chronic inflammation (e.g.
CC rheumatoid arthritis), familial Mediterranean fever (FMF) and
CC systemic amyloidosis found in long-term haemodialysis patients.
XX
XX Sequence 6 AA:
SQ

Query Match 54.3%; Score 25; DB 22; Length 6;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 LIYWA 8
| : | |
Db 2 lvfwa 6

RESULT 11
AAG65040
ID AAG65040 standard; peptide; 7 AA.
XX
XX AAG65040:
AC
XX
XX 09-OCT-2001 (first entry)
DE Human matrix metalloproteinase MMP-8 substrate peptide.
XX
XX Chemical sensor system; microjet; indicator chemistry; ligand analysis;
KW biomedicine; environmental monitoring; biowarefare.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
PT /label= OTHER
PT /note= "modified by Dnp"

FT Modified-site 7
FT /label= OTHER
FT /note= "C-terminal amide"
XX
XX WO200157494-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01553.
XX
XX 20-JAN-2000; 2000US-0177105.
PR 09-NOV-2000; 2000US-0709047.
XX
XX (RESC) UNIV CALIFORNIA.
PA
XX Brown SB, Colston BW, Langry K, Milnovich FP, Simon J, Cox WR;
PI Hayes DJ;
PI WPI: 2001-488913/53.
XX
XX Producing chemical sensors useful for biomedical, environmental,
PT occupational safety, process control, biowarefare applications, by
PT printing indicator chemistries on optically accessible surface by
PT microjet technology -
XX
XX Disclosure; Page 13; 31pp; English.
XX
XX The present invention relates to a method of producing a chemical sensor,
CC involving the use of microjet technology to print one or more indicator
CC chemistries on an optically accessible surface. This is useful for
CC producing a chemical sensor useful for detecting and/or analysing a
CC sample in a fluid or airborne medium, for monitoring hazardous materials
CC such as heavy metals, hydrocarbons and chlorinated hydrocarbons in both
CC ground water and soil of contaminated sites, for making accurate
CC dosimetry measurements of hazardous materials, such as carcinogens or
CC mutagens present in hostile or potentially hostile environments. The
CC sensor systems can be implemented in assembly line type configurations
CC for quality and process control type applications, e.g. measurements of
CC gases emitted from fruits and vegetables and detection of contaminants in
CC soft drink or bottled water solutions. The systems are used for detection
CC of airborne or water-based chemical and biowarefare agents such
CC as anthrax and are suitable for measuring multiple constituents in a
CC small sample volume. The present sequence is a substrate for a matrix
CC metalloproteinase described in the exemplification of the invention.
XX
XX Sequence 7 AA:
SQ

Query Match 54.3%; Score 25; DB 22; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIYWA 8
| : | | |
Db 2 laywa 6

RESULT 12
AAB35978
ID AAB35978 standard; peptide; 7 AA.
XX
XX AAB35978:
AC
XX
XX 01-MAR-2001 (first entry)
DE Collagenase cleavable peptide #7.
XX
XX polymeric drug conjugate; enzymatically cleavable linker; cardiant;
KW antiinflammatory; cytostatic; hepatotropic; neuroprotective; cancer;
KW antibacterial; nephrotropic; immunomodulatory; neoplastic disease;
KW chronic inflammatory disease; acute inflammatory disease;
KW cardiac disease; renal disease; liver disease; lung disease;
KW neurological disease; musculoskeletal disease; immunological disorder;

XX microbial infection.
XX Synthetic.
XX WO200064486-A2.
XX
XX 02-NOV-2000.
XX
XX 28-APR-2000; 2000WO-US11670.
XX
XX 28-APR-1999; 99US-0131404.
XX 02-NOV-1999; 99US-0163090.
XX
XX (VERI-) VERITAS MEDICAL TECHNOLOGIES INC.
XX Pachence JM, Bellinka BA, Ramani T;
XX WPI; 2001-031659/04.
XX
XX Polymeric drug conjugate, for treating diseases associated with organs
PT e.g. liver or heart, has biologically active agents linked to regular
PT repeating linear or branched co-polymers by enzymatically cleavable
PT marker -
XX
XX Claim 43; Page 94; 100pp; English.
XX
XX This invention relates to a polymeric drug conjugate. The drug conjugate
CC comprises biologically active agents conjugated via an enzymatically
CC cleavable linker to a regular repeating linear unit comprising a water
CC soluble polymer segment and a multifunctional chemical moiety, or to a
CC branched polymer comprising two or more water soluble polymer segments
CC each bound to a common multifunctional chemical moiety. The polymeric
CC drug conjugate has antiinflammatory; cytosstatic; cardiatic; hepatotropic;
CC neuroprotective; antibacterial; nephrotropic; and immunomodulatory
CC activity. The drug conjugate is useful for alleviating a pathological
CC condition such as neoplastic diseases, chronic inflammatory diseases
CC acute inflammatory diseases, cardiac diseases, renal diseases, liver
CC diseases, lung diseases, neurological diseases, musculoskeletal diseases
CC and immunological disorders and various microbial infections by
CC modulating immunological or hormonal function. The present sequence
CC represents a peptide which can be used as the enzymatically cleavable
CC linker in the drug conjugate of the invention.
XX
XX Sequence 7 AA;
SQ

Query Match 54.3%; Score 25; DB 22; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 LIYWA 8
| | | |
DB 2 Laywa 6

RESULT 13
ABB56865
ID ABB56865 standard; Peptide; 4 AA.
XX
XX ABB56865;
XX
XX 05-MAR-2002 (first entry)
XX
XX Human SNP related amino acid sequence SEQ ID NO:1430.
XX
XX Human: single nucleotide polymorphism; SNP; polymorphism; cytosstatic;
XX immunosuppressive; antiinflammatory; neuroprotective; antitubercial;
XX autoimmune disease; inflammation; cancer; nervous system disease;
XX infection; polymorphic protein.
XX
XX Homo sapiens.
XX
XX WO200138586-A2.
PN

XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US32311.
XX
XX 24-NOV-1999; 99US-0167383.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Shinkets RA, Leach M;
XX WPI; 2001-355949/37.
XX
XX Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
XX Claim 1; Page 665; 674pp; English.
XX
XX ABB00010 to ABB01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytosstatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterized
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).
XX
XX Sequence 4 AA;
SQ

Query Match 52.2%; Score 24; DB 22; Length 4;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 LIYW 7
| | | |
DB 1 LIYW 4

RESULT 14
AAB82652
ID AAB82652 standard; Peptide; 6 AA.
XX
XX AAB82652;
XX
XX 02-OCT-2001 (first entry)
XX
XX All-D peptide used in Alzheimer's disease vaccine.
XX
XX Alzheimer's disease; amyloidosis; amyloid-related disease;
XX vaccine; therapy; antigen.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..6 /note- "all D-form residues"
XX Modified-site 6 /note- "C-terminal amide"
XX
XX WO200139796-A2.
XX
XX 07-JUN-2001.
PD

XX 29-NOV-2000; 2000WC-CA01413.
 PF
 XX 29-NOV-1999; 99US-0168594.
 PR 28-NOV-2000; 2000US-0724842.
 XX
 PA (NEUR-) NEUROCHEM INC.
 XX
 PI Chalfour R, Hebert L, Kong X, Gervais F;
 XX WPI: 2001-441458/47.
 DR
 XX Preventing/treating amyloid-related disease, especially Alzheimer's
 PT disease, comprises administering antigenic all-D peptide, e.g. as
 PT vaccine, which elicits production of antibodies to prevent
 PT fibrillogenesis and associated cellular toxicity -
 XX
 PS Disclosure: Page 12; 31pp; English.
 XX
 CC The present sequence is that of an all-D peptide suitable for
 CC use for preparing vaccines for preventing or treating Alzheimer's
 CC disease and other amyloid related disorders in humans. It is based
 CC on a portion of amyloid-beta peptide (see AAB82622). Vaccines
 CC are produced using 'non-self' peptides synthesised from the
 CC unnatural D-configuration amino acids to avoid the drawbacks of
 CC 'self' proteins. The all-D peptides need not be aggregated to be
 CC operative or immunogenic. They preferably interact with at
 CC least 1 region of an amyloid protein, e.g. the beta-sheet region
 CC or Gag-binding site region, the amyloid-beta peptide, or their
 CC immunogenic fragments, protein conjugates, immunogenic derivative
 CC peptides and immunogenic peptidomimetics. Examples include all-D
 CC peptides corresponding to residues 1-42, 1-40, 1-35, 1-28, 1-17,
 CC 10-16, 16-21 and 36-42 of the amyloid-beta peptide and the all-D
 CC derivative peptides given in AAB82623-64. The vaccine elicits a
 CC preferential TH-2 or TH-1 response, preventing fibrillogenesis and
 CC associated cellular toxicity. The amyloid related diseases may be
 CC localised amyloidosis, e.g. diabetes type II, neurodegenerative
 CC diseases, e.g. bovine spongiform encephalitis, Creutzfeldt-Jakob
 CC disease, scrapie, cerebral amyloid angiopathy, and prion protein
 CC related disorders, or systemic amyloidosis associated with chronic
 CC infection (e.g. tuberculosis) or chronic inflammation (e.g.
 CC rheumatoid arthritis), familial Mediterranean fever (FMF) and
 CC systemic amyloidosis found in long-term haemodialysis patients.
 XX
 SQ Sequence 6 AA;

Query Match 52.2%; Score 24; DB 22; Length 6;
 Best Local Similarity 60.0%; Pred. No. 6.4e+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 LIYWA 8
 I : : I I
 DB 2 IYWWA 6

RESULT 15
 AAR32362
 ID AAR32362 standard; peptide: 8 AA.
 XX
 AC AAR32362;
 XX
 DT 17-JUN-1993 (first entry)
 XX
 DE Sample peptide to demonstrate a new sequencing method.
 XX
 KM Sequential degradation; perfluoroalkanoic acid, anhydride.
 XX
 OS Synthetic.
 XX
 PN EF529604-A.
 XX
 PD 03-MAR-1993.

XX 26-AUG-1992; 92EP-0114531.
 PF
 XX 28-AUG-1991; 91JP-0217437.
 PR 15-NOV-1991; 91JP-0300818.
 XX
 PA (DASE) SEIKO INSTR INC.
 XX
 PI Satake K, Takamoto K, Tsugita A, Uchida T;
 XX WPI: 1993-068671/09.
 DR
 XX Aminoacid sequencing of proteins or peptide(s) - by sequential
 PT degradation with per:fluoro-alkanoic acid or anhydride
 XX
 PS Example 4; Page 10; 38pp; English.
 XX
 CC The peptide was used to demonstrate a novel method of protein/peptide
 CC sequencing comprising sequentially degrading the peptide from the C
 CC terminus by administering a vapour conty. perfluoroalkanoic acid and
 CC analysing the resultant reaction mixt. The process is simple and avoids
 CC using enzymes or other complicated organic reagents.
 CC See also AAR32360-5.
 XX
 SQ Sequence 8 AA;

Query Match 52.2%; Score 24; DB 14; Length 8;
 Best Local Similarity 66.7%; Pred. No. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKLIY 6
 I : : I : I
 DB 3 phlly 8

Search completed: July 15, 2002, 13:38:45
 Job time: 690 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: July 15, 2002, 13:36:10 ; Search time 12.87 Seconds
(without Alignments)
15.183 Million cell updates/sec

Title: US-09-712-819A-4
Perfect score: 46
Sequence: 1 PKLLIYMA 8

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 48605

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_5/prodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/PCITUS.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	26	56.5	8	US-07-961-307-1	Sequence 1, Appl
2	22	47.8	5	US-08-981-122-11	Sequence 11, Appl
3	22	47.8	6	US-08-180-209-28	Sequence 28, Appl
4	22	47.8	6	US-08-401-512-68	Sequence 68, Appl
5	22	47.8	6	US-08-637-759B-295	Sequence 295, App
6	22	47.8	6	US-08-871-355A-295	Sequence 295, App
7	22	47.8	6	US-09-226-012-103	Sequence 103, App
8	22	47.8	6	US-08-474-853-28	Sequence 28, Appl
9	22	47.8	6	US-09-201-945-295	Sequence 295, App
10	22	47.8	6	PCT-US84-02629-28	Sequence 28, Appl
11	22	47.8	7	US-08-827-962-12	Sequence 12, Appl
12	22	47.8	7	US-08-803-346-73	Sequence 73, Appl
13	22	47.8	8	US-08-594-447-73	Sequence 73, Appl
14	22	47.8	8	US-08-541-964-72	Sequence 72, Appl
15	22	47.8	8	US-08-665-647-87	Sequence 87, Appl
16	21	45.7	5	US-08-068-947-1	Sequence 1, Appl
17	21	45.7	8	US-08-594-447-72	Sequence 72, Appl
18	21	45.7	8	US-08-541-964-71	Sequence 71, Appl
19	21	45.7	8	US-08-621-259A-205	Sequence 205, App
20	21	45.7	8	US-08-665-647-86	Sequence 86, Appl
21	21	45.7	8	US-08-031-538-32	Sequence 32, Appl
22	21	45.7	8	516933-21	Patent No. 516933
23	20	43.5	5	US-08-920-162A-29	Sequence 29, Appl
24	20	43.5	5	US-09-356-931-29	Sequence 29, Appl
25	20	43.5	5	PCT-US84-01321-48	Sequence 48, Appl
26	20	43.5	6	US-09-128-572-4	Sequence 4, Appl
27	20	43.5	6	5208144-9	Patent No. 5208144

28	20	43.5	6	5208144-10	Patent No. 5208144
29	20	43.5	7	US-08-036-210-1	Sequence 1, Appl
30	20	43.5	7	US-08-018-129-2	Sequence 2, Appl
31	20	43.5	7	US-08-018-129-12	Sequence 12, Appl
32	20	43.5	7	US-08-449-609-1	Sequence 1, Appl
33	20	43.5	7	US-08-448-250-2	Sequence 2, Appl
34	20	43.5	7	US-08-448-250-12	Sequence 12, Appl
35	20	43.5	7	US-08-836-480-13	Sequence 13, Appl
36	20	43.5	7	US-08-750-142B-50	Sequence 50, Appl
37	20	43.5	7	US-09-562-897-13	Sequence 13, Appl
38	20	43.5	8	PCT-US94-01321-46	Sequence 46, Appl
39	19	41.3	5	US-08-476-176B-50	Sequence 50, Appl
40	19	41.3	5	US-08-920-162A-18	Sequence 18, Appl
41	19	41.3	5	US-08-920-162A-32	Sequence 32, Appl
42	19	41.3	5	US-08-127-721A-50	Sequence 50, Appl
43	19	41.3	5	US-08-485-246A-50	Sequence 50, Appl
44	19	41.3	5	US-09-356-931-18	Sequence 18, Appl
45	19	41.3	5	US-09-356-931-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-07-961-307-1
; Sequence 1, Application US/07961307
; Patent No. 6090785
; GENERAL INFORMATION:
; APPLICANT: Durette, Philippe L.
; APPLICANT: Esser, Craig K.
; APPLICANT: Hagman, William K.
; APPLICANT: Kopka, Thor E.
; TITLE OF INVENTION: Substituted N-Carboxyalkyl-peptidals
; TITLE OF INVENTION: Derivatives as Antidegenerative Agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 E. Lincoln Avenue, P. O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/961,307
; FILING DATE: 15-OCT-1992
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Rose, David L.
; REGISTRATION NUMBER: 26332
; REFERENCE/DOCKET NUMBER: 18829
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)594-4777
; TELEFAX: (908)594-4720
; TELEX: 138825
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-07-961-307-1
Query Match 56.5%; Score 26; DB 3; Length 8;
Best Local Similarity 57.1%; Pred. No. 1.7e+05;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 PKILLIW 7
DB 2 PKILLIW 8

RESULT 2

US-08-981-122-11
Sequence 11, Application US/08981122B
Patent No. 6127339
GENERAL INFORMATION:
APPLICANT: Hatanaka, Yoshihiro
APPLICANT: Aritomi, Masaharu
TITLE OF INVENTION: Peptide for binding thereto a low density lipoprotein
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/08/981.122B
CURRENT FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: JP 7-176904
PRIOR FILING DATE: 1995-06-21
PRIOR APPLICATION NUMBER: PCT/JP96/01734
PRIOR FILING DATE: 1996-06-21
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 5
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of a peptide synthesized in Example 2 from L-form
Patent No. 6127339
OTHER INFORMATION: F-moc amino acids by solid phase method using a multi-peptide
US-08-981-122-11

Query Match

Best Local Similarity 47.8%; Score 22; DB 3; Length 5;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 IYW 7
DB 3 IYW 5

RESULT 3

US-08-180-209B-28
Sequence 28, Application US/08180209B
Patent No. 5593877
GENERAL INFORMATION:
APPLICANT: King, Te-Plao
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
TITLE OF INVENTION: HALOGENIDASE, AND IMMUNOLOGICAL THERAPIES BASED
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESS: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180.209B
FILING DATE: 11-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/031,400
FILING DATE: 11-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-074 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
US-08-180-209B-28

Query Match

Best Local Similarity 47.8%; Score 22; DB 1; Length 6;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 IYW 7
DB 3 IYW 5

RESULT 4

US-08-401-512-68
Sequence 68, Application US/08401512
Patent No. 5599673
GENERAL INFORMATION:
APPLICANT: Keating, Mark T.
APPLICANT: Curran, Mark E.
APPLICANT: Wang, Qing
TITLE OF INVENTION: Long QT Syndrome Genes
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESS: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3917
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/401.512
FILING DATE: 09-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Stephen A.
REGISTRATION NUMBER: 38,609
REFERENCE/DOCKET NUMBER: 19780-113879
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4848
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-401-512-68

Query Match 47.8%; Score 22; DB 1; Length 6;
Best Local Similarity 40.0%; Pred.No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 LITW 7
| | | | |
DB 1 LIVHW 5

RESULT 5
US-08-637-759B-295
; Sequence 295, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPLMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 295:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-08-637-759B-295

Query Match 47.8%; Score 22; DB 2; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 YWA 8
| | |
DB 3 YWA 5

RESULT 6
US-08-871-355A-295
; Sequence 295, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden

; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPLMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 295:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; HYPOTHEICAL: NO
; US-08-871-355A-295

Query Match 47.8%; Score 22; DB 3; Length 6;
Best Local Similarity 100.0%; Pred.No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 YWA 8
| | |
DB 3 YWA 5

RESULT 7
US-09-226-012-103
; Sequence 103, Application US/09226012
; Patent No. 6207383
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Splawski, Igor
; TITLE OF INVENTION: MUTATIONS IN AND GENOMIC STRUCTURE OF HERG - A LONG QT
; FILE REFERENCE: 2323-136
; CURRENT APPLICATION NUMBER: US/09/226,012
; CURRENT FILING DATE: 1999-01-06
; EARLIER APPLICATION NUMBER: 09/122,847
; EARLIER FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 103
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-226-012-103

Query Match 47.8%; Score 22; DB 4; Length 6;
Best Local Similarity 40.0%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0;

OY 3 LIYW 7
|:::|
DB 1 LIVHW 5

RESULT 8
US-08-474-853-28

Sequence 28; Application US/08474853
Patent No. 6287539

GENERAL INFORMATION:

APPLICANT: King, Te-Piao

TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF

TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,853

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/180,209

FILING DATE: 11-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/031,400

FILING DATE: 11-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

REFERENCE/DOCKET NUMBER: 600-1-074 CIPB

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201 487-5800

TELEFAX: 201 343-1684

TELEX: 133521

INFORMATION FOR SEQ. ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

US-08-474-853-28

Query Match 47.8%; Score 22; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 IYW 7
| | |
DB 3 IYW 5

RESULT 9
US-09-201-945-295

Sequence 295; Application US/09201945
Patent No. 6342215

GENERAL INFORMATION:

APPLICANT: David William Holden

TITLE OF INVENTION: Identification of Genes

NUMBER OF SEQUENCES: 501

CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst

STREET: 2800 One Atlantic Center

STREET: 1201 West Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30309-3450

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/201,945

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/637,759

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: RPMS 101

TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794

TELEFAX: (404) 873-8795

INFORMATION FOR SEQ. ID NO: 295:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

US-09-201-945-295

Query Match 47.8%; Score 22; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 YWA 8
| | |
DB 3 YWA 5

RESULT 10
PCT-US94-02629-28

Sequence 28; Application PC/TUS9402629

GENERAL INFORMATION:

APPLICANT: King, Te-Piao

TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF

TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND

NUMBER OF SEQUENCES: 62

CORRESPONDENCE ADDRESS:

ADDRESSEE: Klauber & Jackson

STREET: 411 Hackensack Avenue

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/02629
FILING DATE: 10-MAR-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/180,209
FILING DATE: 11-JUN-1994
APPLICATION DATA:
APPLICATION NUMBER: US 08/031,400
FILING DATE: 11-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-074 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
PCT-US94-02629-28

Query Match 47.8%; Score 22; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 IYW 7
|||
DB 3 IYW 5

RESULT 11
US-08-827-962-12
Sequence 12, Application US/08827962A
Patent No. 6258944
GENERAL INFORMATION:
APPLICANT: MERCK & CO., INC.
TITLE OF INVENTION: OB RECEPTOR ISOFORMS AND NUCLEIC ACIDS
FILE REFERENCE: 19693
CURRENT APPLICATION NUMBER: US/08/827,962A
CURRENT FILING DATE: 1997-05-06
PRIOR APPLICATION NUMBER: 60/016,899
PRIOR FILING DATE: 1996-05-06
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 7
TYPE: PRT
ORGANISM: Mus Musculus
US-08-827-962-12

Query Match 47.8%; Score 22; DB 4; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 IYMA 8
|||
DB 1 IYWS 4

RESULT 12
US-08-803-346-73
Sequence 73, Application US/08803346
Patent No. 6281346
GENERAL INFORMATION:
APPLICANT: HESS, JOHN W.
APPLICANT: CASKEY, C. THOMAS
APPLICANT: LIU, QINGYUN
APPLICANT: PHILLIPS, MICHAEL SEAN
TITLE OF INVENTION: RAT OB RECEPTORS AND NUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: JOANNE M. GIESSEY - MERCK & CO., INC.
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000
CITY: RAHWAY
STATE: NJ
COUNTRY: USA
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/803,346
FILING DATE: 20-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: GIESSEY, JOANNE M.
REGISTRATION NUMBER: 32,838
REFERENCE/DOCKET NUMBER: 19642Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-3046
TELEFAX: 732-594-4720
TELEX:
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-803-346-73

Query Match 47.8%; Score 22; DB 4; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 IYMA 8
|||
DB 1 IYWS 4

RESULT 13
US-08-594-447-73
Sequence 73, Application US/08594447
Patent No. 5776716
GENERAL INFORMATION:
APPLICANT: Ron, Dorit
APPLICANT: Napolitano, Eugene W.
APPLICANT: Voronova, Anna F.
TITLE OF INVENTION: METHODS FOR IDENTIFYING AGENTS WHICH
BLOCK THE INTERACTION OF PTN WITH PKC-THETA, AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW - Ste. 5500

CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,447
FILING DATE: 31-JAN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22550-20025.24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFORSMH
INFORMATION FOR SEQ ID NO: 73:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..8
OTHER INFORMATION: /label= PRK2-2
US-08-594-447-73

Query Match 47.8%; Score 22; DB 1; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KLIYV 7
: : : : :
Db 1 EISYV 6

RESULT 14
US-08-541-964-72
Sequence 72, Application US/08541964
Patent No. 5783405
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
APPLICANT: Kaurat, Lawrence M.
APPLICANT: Napolitano, Eugene W.
TITLE OF INVENTION: A RAPID SCREENING METHOD FOR EFFECTORS
TITLE OF INVENTION: OF SIGNAL TRANSDUCTION
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW-STE. 5500
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/541,964
FILING DATE: 10-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22550-20025.23
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFORSMH
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..8
OTHER INFORMATION: /label= PRK2-2
US-08-541-964-72

Query Match 47.8%; Score 22; DB 1; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KLIYV 7
: : : : :
Db 1 EISYV 6

RESULT 15
US-08-665-647-87
Sequence 87, Application US/08665647
Patent No. 5935803
GENERAL INFORMATION:
APPLICANT: Dasquez, Nick J.
APPLICANT: Ron, Dorit
APPLICANT: Voronova, Anna F.
APPLICANT: Napolitano, Eugene W.
TITLE OF INVENTION: METHODS TO IDENTIFY IMMUNOMODULATORS
TITLE OF INVENTION: USING COGNATE INTERACTION OF PKC-THETA
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVENUE, NW - Ste. 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/665,647
FILING DATE: 18-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 22550-20025.25
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168
TELEX: 90-4030 MRSNFORSMH
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

```

;
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..8
; OTHER INFORMATION: /label= PRK2-2
US-08-665-647-87
    
```

```

Query Match          47.8%; Score 22; DB 2; Length 8;
Best Local Similarity 33.3%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
    
```

```

QY      2 KILTYW 7
       ::::|
Db       1 EISYVW 6
    
```

```

Search completed: July 15, 2002, 13:39:25
Job time: 195 sec
    
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:01:01 ; Search time 95.45 Seconds
(without alignments)
7.047 Million cell updates/sec

Title: US-09-712-819a-5

Perfect score: 35

Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	34	2 H30607	Ig kappa chain V-I
2	35	100.0	54	2 A25521	Ig kappa chain V-I
3	35	100.0	54	2 J20521	Ig kappa chain V-I
4	35	100.0	71	2 S21526	Ig kappa chain V-I
5	35	100.0	78	2 S34102	Ig kappa chain V-I
6	35	100.0	79	2 S24215	Ig kappa chain V-I
7	35	100.0	81	2 PH1048	Ig light chain V-I
8	35	100.0	82	2 S38560	Ig light chain V-I
9	35	100.0	82	2 S34090	Ig kappa chain V-I
10	35	100.0	83	2 I30607	Ig kappa chain V-I
11	35	100.0	86	2 S16826	Ig kappa chain V-I
12	35	100.0	86	2 S16834	Ig kappa chain V-I
13	35	100.0	86	2 S16836	Ig kappa chain V-I
14	35	100.0	86	2 S34086	Ig kappa chain V-I
15	35	100.0	86	2 S16840	Ig kappa chain V-I
16	35	100.0	86	2 S16837	Ig kappa chain V-I
17	35	100.0	86	2 S16833	Ig kappa chain V-I
18	35	100.0	86	2 S16830	Ig kappa chain V-I
19	35	100.0	86	2 S16824	Ig kappa chain V-I
20	35	100.0	86	2 S16829	Ig kappa chain V-I
21	35	100.0	87	2 S16843	Ig kappa chain V-I
22	35	100.0	87	2 S16842	Ig kappa chain V-I
23	35	100.0	87	2 S16841	Ig kappa chain V-I
24	35	100.0	87	2 S34084	Ig kappa chain V-I
25	35	100.0	87	2 S34083	Ig kappa chain V-I
26	35	100.0	88	2 S21528	Ig kappa chain V-I
27	35	100.0	88	2 S21524	Ig kappa chain V-I
28	35	100.0	88	2 S21525	Ig kappa chain V-I
29	35	100.0	88	2 S21522	Ig kappa chain V-I

30	35	100.0	89	2 S34096	Ig kappa chain V-I
31	35	100.0	90	2 S38561	Ig light chain V-I
32	35	100.0	90	2 I38601	Ig kappa chain V-I
33	35	100.0	91	2 PH1071	Ig light chain V-I
34	35	100.0	91	2 S25462	Ig kappa chain V-I
35	35	100.0	91	2 S37520	Ig kappa chain V-I
36	35	100.0	91	2 S37521	Ig kappa chain V-I
37	35	100.0	91	2 S67940	Ig kappa chain V-I
38	35	100.0	92	2 S37533	Ig kappa chain V-I
39	35	100.0	92	2 S37530	Ig kappa chain V-I
40	35	100.0	92	2 S37529	Ig kappa chain V-I
41	35	100.0	92	2 S37535	Ig kappa chain V-I
42	35	100.0	92	2 S37524	Ig kappa chain V-I
43	35	100.0	92	2 S37513	Ig kappa chain V-I
44	35	100.0	92	2 S37512	Ig kappa chain V-I
45	35	100.0	92	2 S37531	Ig kappa chain V-I

ALIGNMENTS

RESULT 1
H30607
Ig kappa chain V-I region (Bla) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 30-May-1997
C:Accession: H30607
R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; &
J. Immunol. 142, 3158-3163, 1989
A:Title: Structural and idiotypic characterization of the L chains of human IgM autoa
A:Reference number: A30601; M0ID:89215279
A:Accession: H30607
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-34 <CON>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||
DB 23 TDFTLTI 29

RESULT 2
A25521
Ig kappa chain V region (321) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-May-1988 #sequence_revision 21-May-1990 #text_change 09-May-1997
C:Accession: A25521
R:Chen, P.P.; Albrandt, K.; Orida, N.K.; Radoux, V.; Chen, E.Y.; Schrantz, R.; Liu, F
Proc. Natl. Acad. Sci. U.S.A. 83, 8318-8322, 1986
A:Title: Genetic basis for the cross-reactive idiotypes on the light chains of human
A:Reference number: A94135; M0ID:87041448
A:Accession: A25521
A:Molecule type: DNA
A:Residues: 1-54 <CHR>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||
DB 28 TDFTLTI 34

RESULT 3

IG kappa chain V-II region (CPI) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 09-May-1997
C:Accession: J70521
R:Anker, R.; Conley, M.E.; Pollok, B.A.
J. Exp. Med. 169, 2109-2119, 1989
A:Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia
A:Reference number: J70511; MUID:89279157
A:Accession: J70521
A:Molecule type: mRNA
A:Residues: 1-54 <ANK>
A:Note: The sequence shown here is one of eight productive V-D-J mu chain rearrangements
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-49/Domain: V region <VRE>
F:50-54/Domain: J region <JRE>

Query Match 100.0%; Score 35; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
Db 21 TDFTLTI 27

RESULT 4

S21526
IG kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 25-Oct-1996 #text_change 23-Jul-1999
C:Accession: S34082; S21526
R:Magnez, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34082
A:Molecule type: DNA
A>Status: preliminary
A:Residues: 1-71 <MA2>
A:Cross-references: EMBL:X66042; NID:933318; PIDN:CAA6841.1; PID:933319
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
Db 62 TDFTLTI 68

RESULT 5

S34102
IG kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
C:Accession: S34102
R:Magnez, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34102
A:Molecule type: DNA
A>Status: preliminary
A:Residues: 1-78 <MAG>
A:Cross-references: EMBL:X67186

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
Db 60 TDFTLTI 66

RESULT 6

S24215
IG kappa chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24215
R:Shimizu, T.; Iwasato, T.; Yamagishi, H.
J. Exp. Med. 173, 1065-1072, 1991
A:Title: Deletions of immunoglobulin C(kappa) region characterized by the circular ex
A:Reference number: S24214; MUID:91217618
A:Accession: S24215
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-79 <SH1>
A:Cross-references: EMBL:X58202; NID:953718; PIDN:CAA41178.1; PID:9930195
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
Db 42 TDFTLTI 48

RESULT 7

PH1048
Ig light chain V region (clone 165.49) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Jun-1996
C:Accession: PH1048
R:Tillman, D.M.; Jou, N.T.; Hall, R.J.; Marlon, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1048
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-81 <TII>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 100.0%; Score 35; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
Db 54 TDFTLTI 60

RESULT 8

S38560
Ig light chain V region (ASMB1) - mouse (fragment)
C:Species: Mus musculus (house mouse)

```
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S38560
R:Monestier, M.; Losman, L.J.; Novick, K.E.; Aris, J.P.
submitted to the EMBL Data Library, September 1993
A:Description: Molecular analysis of mercury-induced anti-nucleolar antibodies in H-2s M
A:Reference number: S38559
A:Accession: S38560
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-82 <MON>
A:Cross-references: EMBL:X75102; NID:q414147; PIDN:CAAS2993.1; PID:q414148
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match          100.0%; Score 35; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

RESULT 9
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34090
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A>Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34090
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-82 <WAG>
A:Cross-references: EMBL:X67174
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMW>

Query Match          100.0%; Score 35; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

RESULT 10
Ig kappa chain V-III region (Wei) - human (fragments)
C:Species: Homo sapiens (man)
C>Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 09-May-1997
C:Accession: S30607
R:Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjona, M.L.; Fernandez, J.; Carson, D.; Sold
J. Immunol. 142, 3158-3163, 1989
A>Title: Structural and idiotype characterization of the L chains of human Igm autoanti
A:Reference number: A30601; MUID:9215279
A:Accession: S30607
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-83 <GON>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match          100.0%; Score 35; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 70 TDFTLTI 76

RESULT 11
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16826; S34101
R:Balsan, G.; Kuntz, J.L.; Pasquali, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A>Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid
A:Reference number: S16823; MUID:91243737
A:Accession: S16826
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54824; NID:q33653; PIDN:CAAS8593.1; PID:q33654
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A>Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distrib
A:Reference number: S34076; MUID:93170387
A:Accession: S34101
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <WAG>
A:Cross-references: EMBL:X67185
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMW>

Query Match          100.0%; Score 35; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

RESULT 12
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16834
R:Balsan, G.; Kuntz, J.L.; Pasquali, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A>Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid
A:Reference number: S16823; MUID:91243737
A:Accession: S16834
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54832
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMW>

Query Match          100.0%; Score 35; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66
```

```

RESULT 13
S16836
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16836
R:Blaisson, G.; Kuntz, J.L.; Pasquall, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid fac
A:Reference number: S16823; MUID:91243737
A:Accession: S16836
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54834
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:7-81/Domain: Immunoglobulin homology <IMM>

```

```

Query Match          100.0%; Score 35; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

```

```

RESULT 14
S34086
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34086
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34086
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <WAG>
A:Cross-references: EMBL:X67169
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:9-83/Domain: Immunoglobulin homology <IMM>

```

```

Query Match          100.0%; Score 35; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TDFTLTI 7
    |||||
Db 62 TDFTLTI 68

```

```

RESULT 15
S16840
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16840
R:Blaisson, G.; Kuntz, J.L.; Pasquall, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid fac
A:Reference number: S16823; MUID:91243737
A:Accession: S16840
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>

```

```

A:Cross-references: EMBL:X54838
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:7-81/Domain: Immunoglobulin homology <IMM>

```

```

Query Match          100.0%; Score 35; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TDFTLTI 7
    |||||
Db 60 TDFTLTI 66

```

```

Search completed: July 15, 2002, 13:01:01
Job time: 467 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:23:32 ; Search time 44.9 Seconds
(without alignments)
6.036 Million cell updates/sec

Title: US-09-712-819A-5

Perfect score: 35

Sequence: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	100	KV3C_HUMAN	P01621 homo sapien
2	35	100.0	107	KV04_RABIT	P01685 oryctolagus
3	35	100.0	107	KV1D_HUMAN	P01596 homo sapien
4	35	100.0	108	KV06_RABIT	P01687 oryctolagus
5	35	100.0	108	KV1E_HUMAN	P01597 homo sapien
6	35	100.0	108	KV1H_HUMAN	P01600 homo sapien
7	35	100.0	108	KV1Q_HUMAN	P01609 homo sapien
8	35	100.0	108	KV5P_MOUSE	P01649 mus musculu
9	35	100.0	108	KV5Q_MOUSE	P01650 mus musculu
10	35	100.0	108	KV5R_MOUSE	P01651 mus musculu
11	35	100.0	108	KV5S_MOUSE	P01652 mus musculu
12	35	100.0	108	KV5T_MOUSE	P01653 mus musculu
13	35	100.0	109	KV01_RAT	P01681 rattus norv
14	35	100.0	109	KV1T_HUMAN	P01612 homo sapien
15	35	100.0	109	KV3B_HUMAN	P01620 homo sapien
16	35	100.0	109	KV3D_HUMAN	P01622 homo sapien
17	35	100.0	109	KV3E_HUMAN	P01623 homo sapien
18	35	100.0	109	KV3G_HUMAN	P01626 homo sapien
19	35	100.0	109	KV3H_HUMAN	P01660 mus musculu
20	35	100.0	111	KV3J_MOUSE	P01662 mus musculu
21	35	100.0	111	KV3K_MOUSE	P01663 mus musculu
22	35	100.0	114	KV1A_MOUSE	P01632 mus musculu
23	35	100.0	114	KV4A_HUMAN	P01625 homo sapien
24	35	100.0	115	KV3I_HUMAN	P01643 homo sapien
25	35	100.0	115	KV5A_HUMAN	P06315 homo sapien
26	35	100.0	116	KV3J_HUMAN	P04443 homo sapien
27	35	100.0	117	KV1I_HUMAN	P01601 homo sapien
28	35	100.0	121	KV4O_HUMAN	P06312 homo sapien
29	35	100.0	129	KV1M_HUMAN	P04431 homo sapien
30	35	100.0	129	KV3L_HUMAN	P18135 homo sapien
31	35	100.0	129	KV3M_HUMAN	P18136 homo sapien
32	35	100.0	131	KV3I_MOUSE	P01661 mus musculu
33	35	100.0	133	KV4B_HUMAN	P06313 homo sapien

34	35	100.0	134	1	KV4C_HUMAN	P06314 homo sapien
35	35	100.0	136	1	KV5B_MOUSE	P01634 mus musculu
36	33	94.3	104	1	KV17_RABIT	P01698 oryctolagus
37	33	94.3	112	1	KV3C_MOUSE	P01658 mus musculu
38	33	94.3	132	1	KV3F_MOUSE	P01658 mus musculu
39	32	91.4	285	1	ALXM_PHOS4	P39049 photodacter
40	31	88.6	108	1	KV05_RABIT	P01686 oryctolagus
41	31	88.6	108	1	KV07_RABIT	P01688 oryctolagus
42	31	88.6	108	1	KV08_RABIT	P01689 oryctolagus
43	31	88.6	108	1	KV1A_HUMAN	P01593 homo sapien
44	31	88.6	108	1	KV1E_HUMAN	P01598 homo sapien
45	31	88.6	108	1	KV1G_HUMAN	P01599 homo sapien

ALIGNMENTS

```

RESULT 1
KV3C_HUMAN          STANDARD;          PRT;          100 AA.
ID   KV3C_HUMAN
AC   P01621:
DT   21-JUL-1986 (Rel. 01, Created)
DI   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V-II region NC9 precursor (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84093600; PubMed=6419127;
RA   Bentley D.L.;
RT   "Most kappa immunoglobulin mRNA in human lymphocytes is homologous to
    a small family of germ-line V genes.";
RL   Nature 307:77-80(1984).
CC   -I- MISCELLANEOUS; THIS GENE WAS ISOLATED FROM THE NC9/9.1 HYBRIDOMA.
DR   PIR; A01894; K3HNG.
DR   HSSP; P01789; IMCP.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_V.
DR   Pfam; PF00047; Ig_1.
KW   SMART; SM00406; IgV_1.
KW   Immunoglobulin V region; Signal; HydrIdoma.
FT   NON_TER          1
FT   SIGNAL           1
FT   CHAIN            5
FT   DISULFID         27
FT   NON_TER          100
SQ   SEQUENCE      100 AA; 10729 MW; 5D9AF363CC52632F CRC64;
Query Match          100.0%; Score 35; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY   1 TDFTLTI 7
Db    74 TDFTLTI 80
RESULT 2
KV04_RABIT
ID   KV04_RABIT          STANDARD;          PRT;          107 AA.
AC   P01685:
DT   21-JUL-1986 (Rel. 01, Created)
DI   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig kappa chain V region 4135.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;

```

RN [1]
RP SEQUENCE:
RX MEDLINE-75133568; PubMed-1091650;
RA Chen K.C.S., Kindt T.J., Krause R.M.;
RT "Primary structure of the L chain from a rabbit homogeneous antibody
RT to streptococcal carbohydrate, II. Sequence determination of peptides
RT from tryptic and peptic digests";
RL J. Biol. Chem. 250:3289-3296(1975).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE B4-TYPE C REGION IS ALSO GIVEN.
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM ANTIBODY TO THE
CC SPECIFIC CARBOHYDRATE OF GROUP C STREPTOCOCCI AND WAS ISOLATED
CC FROM THE SERUM OF A SINGLE RABBIT.
DR HSSP: P80362; 1MTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 24 FRAMEWORK-1.
FT DOMAIN 25 35 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 36 50 FRAMEWORK-2.
FT DOMAIN 51 57 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 58 89 FRAMEWORK-3.
FT DOMAIN 90 96 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 97 106 FRAMEWORK-4.
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11182 MW; 8F84C5FEE60B7222 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFLTIT 7
DB 70 TDFLTIT 76

RESULT 3
KVLD_HUMAN STANDARD; PRT; 107 AA.
ID P01597;
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DY 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region CAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE:
RX MEDLINE-75075135; PubMed-4216454;
RA Milstein C.P., Deverson E.V.;
RT "Primary structure of kappa light chain from a human myeloma
RT protein";
RL Eur. J. Biochem. 49:377-391(1974).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
DR HSSP: P80362; 1MTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; glycoprotein.
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .).
FT NON_TER 107 107
SQ SEQUENCE 107 AA; 11703 MW; E1BFD0F9844C346 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TDFLTIT 7
DB 69 TDFLTIT 75

RESULT 4
KV06_RABIT STANDARD; PRT; 108 AA.
ID P01687;
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DY 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V region BS-5.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE:
RX MEDLINE-75127348; PubMed-4156171;
RA Jacon J.-C.;
RT "Comparison of the amino acid sequences of the variable regions of
RT light chains derived from two homogeneous rabbit anti-pneumococcal
RT antibodies";
RL Biochem. J. 141:15-25(1974).
RN [2]
RA AMIDES AT POSITIONS 37 AND 89.
RA Jacon J.-C.;
RT Submitted (JUN-1975) to the PIR data bank.
CC -1- MISCELLANEOUS: THIS CHAIN DIFFERS FROM THE KAPPA CHAIN FROM RABBIT
CC BS-1, ALSO OBTAINED FROM ANTIBODY TO TYPE III PNEUMOCOCCI, AT 8
CC POSITIONS IN THE V REGION.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 107 FRAMEWORK-4.
FT DISULFID 23 88
FT DISULFID 80 80 WITH A CYS IN THE C REGION.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11281 MW; CFB6D10DBB7A5FCE CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFLTIT 7
DB 69 TDFLTIT 75

RESULT 5
KVLE_HUMAN STANDARD; PRT; 108 AA.
ID P01597;
AC 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DY 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region DEE.


```
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Ig kappa chain V-V regions (Anti-arsenate antibodies).
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NM NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC STRAIN=A/J;
RX MEDLINE=77250895; PubMed=70482;
RA Capra J.D., Tung A.S., Nisicoff A.;
RT "Structural studies on induced antibodies with defined idiotypic
RT specificities. V. The complete amino acid sequence of the light chain
RT variable regions of anti-P-azophenylarsenate antibodies from A/J mice
RT bearing a cross-reactive idiotype."
RL J. Immunol. 119:993-999(1977).
CC -1- MISCELLANEOUS: THE MIXTURE SEQUENCED CONTAINED AT LEAST TWO OR
CC THREE DIFFERENT LIGHT CHAINS. PEPTIDES CONTAINING THE FOLLOWING
CC SUBSTITUTIONS WERE ALSO ISOLATED: 3-VAL, 10-ILE, AND 12-LEU;
CC 23-LEU: 36-PHE, 41-GLU, AND 43-ALA; 63-THR AND 68-ARG; 76-SER,
CC 77-SER, 78-VAL, 80-ALA, AND 85-ASP; AND 100-GLN AND 107-ARG.
DR PIR: A01928; KWSAA.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_V.
KW SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Antisensate antibody.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 3 49 FRAMEWORK-2.
FT DOMAIN 5 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 12056 MW; AE2861B6AAC09DD2 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 9
KV5Q_MOUSE STANDARD; PRT; 108 AA.
ID KV5Q_MOUSE
AC P01650;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region UPC 61.
DE Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudikoff S., Potter M.;
RT "The structural basis of a hapten-inhibitable kappa-chain idiotype."
RL J. Immunol. 122:1905-1910(1979).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: A01929; KWSG1.
DR HSSP: P80362; IWTU.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_V.
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;
```

```
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 3 49 FRAMEWORK-2.
FT DOMAIN 5 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 FRAMEWORK-4.
FT DOMAIN 98 108 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11809 MW; FAED4A36076F2AFE CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 10
KV5R_MOUSE STANDARD; PRT; 108 AA.
ID KV5R_MOUSE
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
DE Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudikoff S., Potter M.;
RT "The structural basis of a hapten-inhibitable kappa-chain idiotype."
RL J. Immunol. 122:1905-1910(1979).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B92808; KWS09.
DR HSSP: P01607; IRET.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_V.
KW SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 34 FRAMEWORK-2.
FT DOMAIN 3 49 FRAMEWORK-3.
FT DOMAIN 5 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-4.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 11
KV5S_MOUSE
```

```

ID  KV5_MOUSE  STANDARD;  PRT;  108 AA.
AC  P01652;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DE  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Ig kappa chain V-V region J606.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=82099361; PubMed=6798111;
RA  Johnson N., Stankard J., Paul L., Hood L.;
RT  "The complete V domain amino acid sequences of two myeloma Inulin-
binding proteins."
RL  J. Immunol. 128:302-307(1982).
CC  -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC  BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR  HSSP; A92811; KMS06.
DR  InterPro; IPR003006; IG_MHC.
DR  InterPro; IPR003596; IG_V.
DR  Pfam; PF00047; Ig; 1.
DR  SMART; SM00406; IGV; 1.
KW  Immunoglobulin V region.
FT  DOMAIN 1 23
FT  DOMAIN 2 34
FT  DOMAIN 3 49
FT  DOMAIN 4 56
FT  DOMAIN 5 57
FT  DOMAIN 6 88
FT  DOMAIN 7 89
FT  DOMAIN 8 97
FT  DOMAIN 9 108
FT  DISULFID 23 88
FT  NON_TER 108
SQ  SEQUENCE 108 AA; 11810 MW; 8DE4DD31076F2AFB CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
DB 69 TDFTLTI 75

RESULT 12
KV5T_MOUSE
ID KV5T_MOUSE STANDARD; PRT; 108 AA.
AC P01653;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region W3082.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma Inulin-
binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR; B92811; KMS82.
DR HSSP; P80362; IGMT.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.

```

```

DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 2 34
FT DOMAIN 3 49
FT DOMAIN 4 56
FT DOMAIN 5 57
FT DOMAIN 6 88
FT DOMAIN 7 89
FT DOMAIN 8 97
FT DOMAIN 9 108
FT DISULFID 23 88
FT NON_TER 108
SQ SEQUENCE 108 AA; 11850 MW; C5C145DC376F30CD CRC64;

Query Match 100.0%; Score 35; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
DB 69 TDFTLTI 75

RESULT 13
KV01_RAT
ID KV01_RAT STANDARD; PRT; 109 AA.
AC P01681;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V region S211.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX STRAIN=10J;
RX MEDLINE=7521238; PubMed=807630;
RA Starace V., Querinjean P.;
RT "The primary structure of a rat kappa Bence Jones protein:
phylogenetic relationships of V- and C-region genes."
RL J. Immunol. 115:59-62(1975).
CC -I- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
DR PIR; A01944; KVR21.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23
FT DOMAIN 2 35
FT DOMAIN 3 50
FT DOMAIN 4 57
FT DOMAIN 5 58
FT DOMAIN 6 89
FT DOMAIN 7 90
FT DOMAIN 8 98
FT DOMAIN 9 108
FT NON_TER 109
SQ SEQUENCE 109 AA; 11947 MW; A25BFB99D5CB4C6 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
DB 70 TDFTLTI 76

RESULT 14
KV1T_HUMAN
ID KV1T_HUMAN STANDARD; PRT; 109 AA.

```

AC P01612;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-I region Mkv.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=83081018; PubMed=6816713;
 RA Eullitz M., Linke R.P.;
 RT Primary structure of the variable part of an amyloidogenic
 RT Bence-Jones Protein (Mkv). An unusual insertion in the third
 RT hypervariable region of a human kappa-immunoglobulin light chain.;
 RL Hoppe-Seivler's 2. Physiol. Chem. 363:1347-1358(1982).
 CC -1- MISCELLANEOUS: ANOTHER FORM THAT LACKED RESIDUES 1-3 WAS ALSO
 CC FOUND.
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
 DR PIR: A01879; KIHUV.
 DR HSSP: P01730; IMIO.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_V.
 DR SMART: SM00406; IgV; 1.
 KW Immunoglobulin V region.
 FT DOMAIN 1 23 FRAMEWORK-1.
 FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 35 49 FRAMEWORK-2.
 FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 57 66 FRAMEWORK-3.
 FT DOMAIN 89 98 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 99 108 FRAMEWORK-4.
 FT DISULFID 23 88 BY SIMILARITY.
 FT NON_TER 109 109
 SQ SEQUENCE 109 AA; 11870 MW; B6ABF451SD5F5A0 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 |||||
 DB 69 TDFTLTI 75

RESULT 15
 KV3B_HUMAN STANDARD; PRT; 109 AA.
 AC P01620;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-III region SIE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=82046598; PubMed=6794615;
 RA Andrews D.W., Capra J.D.;
 RT Amino acid sequence of the variable regions of light chains from two
 RT idiotypically cross-reactive human Igm anti-gamma-globulins of the Wa
 RT group.;
 RL Biochemistry 20:5816-5822(1981).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
 CC GLOBULIN ACTIVITY.
 DR PIR: A01892; K3HUSI.
 DR HSSP: P01789; IMCP.
 DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig_V.
 DR SMART: SM00406; IgV; 1.
 KW Immunoglobulin V region.
 FT DISULFID 23 89 BY SIMILARITY.
 FT NON_TER 109 109
 SQ SEQUENCE 109 AA; 11775 MW; 7689C3ECD646FFB4 CRC64;

Query Match 100.0%; Score 35; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 0.35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 |||||
 DB 70 TDFTLTI 76

Search completed: July 15, 2002, 13:23:32
 Job time: 1448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      July 15, 2002, 13:22:44 ; Search time 172.49 Seconds
              (without alignments)
              7.020 Million cell updates/sec
```

Title:	US-09-712-819A-5
Perfect score:	35
Sequence:	1 TDFTLTI 7

Scoring table: BLOSUM62

```
Searched:      562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters:  5622222
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
```

```
Database :
    1:  sp.archaea:*
    2:  sp.bacteria:*
    3:  sp.fungi:*
    4:  sp.human:*
    5:  sp.invertebrate:*
    6:  sp.mammal:*
    7:  sp.mmc:*
    8:  sp.organelle:*
    9:  sp.phage:*
   10:  sp.plant:*
   11:  sp.podent:*
   12:  sp.virus:*
   13:  sp.vertebrate:*
   14:  sp.unclassified:*
   15:  sp.virus:*
   16:  sp.bacteriap:*
   17:  sp.archaeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	100.0	107	4	Q9UL81	Q9UL81 homo sapien
2	35	100.0	107	4	Q96SA9	Q96SA9 homo sapien
3	35	100.0	107	11	Q9ERZ9	Q96SA9 homo sapien
4	35	100.0	108	4	Q9UL79	Q9ERZ9 mus muscul
5	35	100.0	108	4	Q9UL77	Q9UL79 homo sapien
6	35	100.0	108	4	Q9UL70	Q9UL77 homo sapien
7	35	100.0	109	4	Q9UL86	Q9UL70 homo sapien
8	35	100.0	109	4	Q9UL78	Q9UL86 homo sapien
9	32	91.4	109	6	Q9N0W5	Q9UL78 homo sapien
10	32	91.4	557	16	Q99XR2	Q9N0W5 oryctolagus
11	31	88.6	99	11	Q9JL74	Q99XR2 streptococ
12	31	88.6	108	4	Q9UL83	Q9JL74 mus muscul
13	31	88.6	109	4	Q9UL85	Q9UL83 homo sapien
14	31	88.6	110	9	Q94MK5	Q9UL85 homo sapien
15	31	88.6	110	9	Q94MK3	Q94MK5 bacterioph
16	31	88.6	110	9	Q94MK2	Q94MK3 bacterioph
						Q94MK2 bacterioph

17	31	88.6	110	9	Q94MC9	Q94mc1 bacterioph
18	31	88.6	110	9	Q94MB9	Q94mb9 bacterioph
19	31	88.6	110	9	Q94LZ7	Q94Lz7 bacterioph
20	31	88.6	110	9	Q94LW6	Q94Lw6 bacterioph
21	31	88.6	110	9	Q94LY9	Q94Ly9 bacterioph
22	31	88.6	572	10	Q94RZ7	Q94rz7 phytophthor
23	31	88.6	16	Q94RZ7	P74440 synchocyst	
24	30	85.7	110	9	Q94MC4	Q94mc4 bacterioph
25	30	85.7	111	11	Q920E9	Q920e9 mus musculu
26	30	85.7	229	9	Q94JH8	Q94jh8 pseudomonas
27	30	85.7	276	8	Q02656	Q02656 podospora
28	30	85.7	280	5	Q17867	Q17867 caenorhabdi
29	30	85.7	320	11	Q99LQ3	Q99lq3 mus musculu
30	30	85.7	412	5	Q45223	Q45223 caenorhabdi
31	30	85.7	415	5	Q97ZG2	Q97zgz caenorhabdi
32	30	85.7	417	4	Q94VQ9	Q94vq9 homo sapien
33	30	85.7	695	16	Q93CD7	Q93cd7 pasteurella
34	30	85.7	886	11	Q91WC9	Q91w9 mus musculu
35	30	85.7	903	6	Q96LY5	Q96ly5 oryctolagus
36	30	85.7	1865	2	Q9XCJ1	Q9xcj1 salmonella
37	29	82.9	104	11	Q93J82	Q93j82 mus musculu
38	29	82.9	114	4	Q90L80	Q90l80 homo sapien
39	29	82.9	238	11	Q99M37	Q99m37 mus musculu
40	29	82.9	241	16	Q92NM7	Q92nm7 rhizobium m
41	29	82.9	247	9	Q94IE6	Q94ie6 lactobacilli
42	29	82.9	269	5	Q008I4	Q008i4 tritriclomo
43	29	82.9	307	3	Q74782	Q74782 schizosacch
44	29	82.9	351	17	Q26846	Q26846 methanother
45	29	82.9	394	16	Q35679	Q35679 synchocyst

ALIGNMENTS

RESULT	1
09UL81	
ID	09UL81
AC	PRELIMINARY:
AC	PRT: 107 AA.
DT	01-MAY-2000 (Tremblrel, 13, Created)
DT	01-MAY-2000 (Tremblrel, 13, Last sequence update)
DT	01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE	MOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE	(FRAGMENT).
OS	<i>Homo sapiens (Human)</i> .
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98277139; PubMed=9614934;
RA	Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Betney S.M.,
RA	Young D.C.;
RT	"Myosin-reactive autoantibodies in rheumatic fever and normal
RT	fetus.";
RL	Clin. Immunol. Immunopathol. 87:184-192(1998).
DR	EMBL: AF035033; AAD56269.1; -
DR	HSSP: P01607; 1RET
DR	InterPro: IPR003006; Ig_MHC.
DR	InterPro: IPR003596; Ig_V.
DR	Pfam: PF00047; Ig_1.
DR	SMART: SM00406; IgV_1.
FT	NON_TER 1
FT	NON_TER 107
SO	SEQUENCE 107 AA; 11501 MW; 070549FDE0754748 CRC64;

Query Match	100.0%;	Score 35;	DB 4;	Length 107;
Best Local Similarity	100.0%;	Pred. No. 3;		
Matches	7;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	TDFTLTI	7
Db	69	TDFTLTI	75

```

RESULT 2
O96SA9 PRELIMINARY; PRT; 107 AA.
AC O96SA9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ANTI-STREPTOCOCCAL/ANTI-MYOSIN IMMUNOGLOBULIN KAPPA LIGHT CHAIN
DE VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98375893; PubMed=9712075;
RA Adderson E.E., Shikhan A.R., Ward K.E., Cunningham M.W.;
RT "Molecular analysis of polyclonal monoclonal antibodies from
RT rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin
RT antibody V region genes."
RL J. Immunol. 161:2020-2031(1998).
DR EMBL; U96396; AAB68785.1; -.
FT NON_TER 1
FT NON_TER 1
FT SEQUENCE 107 AA; 11520 MW; 4BB43E9C5B577F16 CRC64;

Query Match 100.0%; Score 35; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 3
O9ER29 PRELIMINARY; PRT; 107 AA.
AC O9ER29;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ANTI HUMAN TNF-ALPHA LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
RA "Cloning and sequencing of the light chain fragment of variable region
RT genes of an anti-tTNF- $\alpha$  monoclonal antibody."
RL J. Cell. Mol. Immunol. 12:21-26(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RA "Construction and sequencing of the single-chain antibody gene of a
RT human TNF- $\alpha$  specific monoclonal antibody."
RL T1 4 Chun 1 Ta Hsueh Hsueh Pao 19:373-376(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
RA Submitted (May-2000) to the EMBL/GenBank/DBD databases.
DR EMBL; AF262753; AAG3804.1; -.
DR HSSP; P80362; 1MTL.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00409; IG_1.
DR SMART; SM00409; IG_1.

```

```

DR SMART; SM00406; IGV_1.
FT NON_TER 107
FT NON_TER 107
SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

Query Match 100.0%; Score 35; DB 11; Length 107;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TDFTLTI 7
Db 72 TDFTLTI 78

RESULT 4
O9UL79 PRELIMINARY; PRT; 108 AA.
AC O9UL79;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035035; AAD56271.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; IG_1.
DR SMART; SM00406; IGV_1.
FT NON_TER 1
FT NON_TER 1
FT SEQUENCE 108 AA; 11787 MW; DB5845E19724FB4E CRC64;

Query Match 100.0%; Score 35; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 5
O9UL77 PRELIMINARY; PRT; 108 AA.
AC O9UL77;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;

```



```

RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035037; AAD56273.1; -.
DR HSSP: P01607; 1REI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
RT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11738 MW; C06681716C4D16F3 CRC64;

Query Match 100.0%; Score 35; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 6
Q9UL70 PRELIMINARY; PRT; 108 AA.
AC Q9UL70;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035044; AAD56280.1; -.
DR HSSP: P01607; 1REI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
RT NON_TER 1
FT NON_TER 108
SQ SEQUENCE 108 AA; 11633 MW; B7BEDC3E41FCCA37 CRC64;

Query Match 100.0%; Score 35; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 7
Q9UL86 PRELIMINARY; PRT; 109 AA.
AC Q9UL86;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN KAPPA CHAIN VARIABLE REGION
DE (FRAGMENT).

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035028; AAD56264.1; -.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
RT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11928 MW; 243325F72C7DAC83 CRC64;

Query Match 100.0%; Score 35; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 70 TDFTLTI 76

RESULT 8
Q9UL78 PRELIMINARY; PRT; 109 AA.
AC Q9UL78;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035036; AAD56272.1; -.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; Igv; 1.
RT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11646 MW; 5F675C52EC7EE197 CRC64;

Query Match 100.0%; Score 35; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 70 TDFTLTI 76

```

RESULT 9
Q9N0M5 PRELIMINARY; PRT: 109 AA.
AC Q9N0M5:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ANTI-HUMAN A33 LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20250927; PubMed=10788485;
RA Rader C., Rittler G., Nachan S., Elia M., Jungbluth A.A.,
RA Cohen L.S., Welt S., Old L.J., Barbas C.F. III.;
RT "The rabbit antibody repertoire as a novel source for the generation
RT of therapeutic human antibodies."
RL J. Biol. Chem. 275:13668-13676(2000).
DR EMBL: AF245502; AAF68449.1; -.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; 1g; 1.
DR SMART: SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 109
SQ SEQUENCE 109 AA; 11333 MW; BD8B396E75F94FB CRC64;

Query Match 91.4%; Score 32; DB 6; Length 109;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TDFTLTI 7
Db 69 TDFTLTI 75

RESULT 10
Q99XR2 PRELIMINARY; PRT: 557 AA.
AC Q99XR2:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE FORMATE-TETRAHYDROFOLATE LIGASE (EC 6.3.4.3).
GN FHS.2 OR SPY2085.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lal H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL: AE006628; AAK34738.1; -.
DR InterPro: IPR000559; FTHFS.
DR Pfam: PF01268; FTHFS; 1.
DR PROSITE: PS00722; FTHFS_2; 1.
KM Llgase; Complete proteome.
SQ SEQUENCE 557 AA; 59053 MW; CB07C9FCE90B34E7 CRC64;

Query Match 91.4%; Score 32; DB 16; Length 557;
Best Local Similarity 85.7%; Pred. No. 75;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TDFTLTI 7
Db 505 TDFTLTI 511

RESULT 11
Q9JL74 PRELIMINARY; PRT: 99 AA.
AC Q9JL74:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ANTI-MYOSIN IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BA1B/C;
RX MEDLINE=20448942; PubMed=10992488;
RA Malkiel S., Liao L., Cunningham M.W., Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of
RT streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
RT with cardiac myosin."
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL: AF206032; AAF69330.1; -.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; 1g; 1.
DR SMART: SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 10939 MW; 3B25D0E78453324 CRC64;

Query Match 88.6%; Score 31; DB 11; Length 99;
Best Local Similarity 85.7%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
Db 61 TDFTLTI 67

RESULT 12
Q9UL83 PRELIMINARY; PRT: 108 AA.
AC Q9UL83:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RX Wu X., Liu B., Van der Merwe P.L., Kalls N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035031; AAD56267.1; -.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.

DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 FT NON_TER 1
 SQ SEQUENCE 108 AA; 11834 MW; 9F9C5A92EBA96EEA CRC64;

Query Match 88.6%; Score 31; DB 4; Length 108;
 Best Local Similarity 85.7%; Pred. No. 23;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 1:|||||
 DB 69 TDFTLTI 75

RESULT 13
 O9UL85 PRELIMINARY; PRT; 109 AA.
 AC O9UL85:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN KAPPA CHAIN VARIABLE REGION (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035029; AAD56265.1; -.
 DR HSP; P80362; IRTL.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 FT NON_TER 1
 SQ SEQUENCE 109 AA; 11761 MW; FBIE43E7C7AFACCC CRC64;

Query Match 88.6%; Score 31; DB 4; Length 109;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 1:|||||
 DB 69 TDFTLTI 75

RESULT 14
 O94MG5 PRELIMINARY; PRT; 110 AA.
 AC O94MG5:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROTEIN GP30.8.
 GN 30.8.
 OS Bacteriophage K3.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
 OX NCBI_TaxID=10674;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Truncate L., Nivinskas R.;

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ315751; CAC42994.1; -.
 SQ SEQUENCE 110 AA; 12892 MW; 1685D7E297D853F7 CRC64;

Query Match 88.6%; Score 31; DB 9; Length 110;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 1:|||||
 DB 44 TDFTLTI 50

RESULT 15
 O94MC3 PRELIMINARY; PRT; 110 AA.
 AC O94MC3:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROTEIN GP30.8.
 GN 30.8.
 OS Bacteriophage Pol.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
 OX NCBI_TaxID=36341;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Truncate L., Nivinskas R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ315760; CAC43000.1; -.
 SQ SEQUENCE 110 AA; 12953 MW; B265DD59971389B6 CRC64;

Query Match 88.6%; Score 31; DB 9; Length 110;
 Best Local Similarity 85.7%; Pred. No. 24;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 1:|||||
 DB 44 TDFTLTI 50

Search completed: July 15, 2002, 13:22:45
 Job time: 1481 sec

10

11

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:57:56 ; Search time 228.39 Seconds
(without alignments)
3.404 Million cell updates/sec

Title: US-09-712-819a-5

Perfect score: 35

Sequence: 1 TDFTLRT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Genesec_032802:*

1: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	16	20	AAV41874
2	35	100.0	16	22	AAU25389
3	35	100.0	16	22	AAU15733
4	35	100.0	32	17	AA887043
5	35	100.0	32	20	AAV52745
6	35	100.0	32	22	AA898286
7	35	100.0	32	22	AA898288
8	35	100.0	32	22	AA897666
9	35	100.0	32	22	AA897668
10	35	100.0	74	19	AAW62805
11	35	100.0	76	20	AAW80961

12	35	100.0	82	19	AAW62807	Amino acid sequenc
13	35	100.0	84	17	AAW14491	Monoclonal antibod
14	35	100.0	84	17	AA899878	Monoclonal antibod
15	35	100.0	86	19	AAW62805	Amino acid sequenc
16	35	100.0	88	19	AAW59619	Anti-RSV F protein
17	35	100.0	88	21	AAV56651	Partial peptide fr
18	35	100.0	88	21	AAV56653	Partial peptide fr
19	35	100.0	88	21	AAV56654	Partial peptide fr
20	35	100.0	88	21	AAV56655	Partial peptide fr
21	35	100.0	88	21	AAV56657	Partial peptide fr
22	35	100.0	88	21	AAV56659	Partial peptide fr
23	35	100.0	88	21	AAV56668	Partial peptide fr
24	35	100.0	91	20	AAW95478	Mouse derived RT3
25	35	100.0	93	15	AAW54321	Anti-HIV gp120 Imm
26	35	100.0	93	17	AAW01290	VL region of HIV n
27	35	100.0	93	21	AAV95142	Anti-gp120 antibod
28	35	100.0	93	21	AAV98251	Anti-gp120 antibod
29	35	100.0	94	19	AAW62808	Amino acid sequenc
30	35	100.0	94	21	AAW56669	Partial peptide fr
31	35	100.0	94	22	AAV56669	Partial peptide fr
32	35	100.0	94	22	AAV56670	Partial peptide fr
33	35	100.0	95	16	AAW72061	Partial peptide fr
34	35	100.0	95	16	AAW72062	Partial peptide fr
35	35	100.0	95	16	AAW72063	Partial peptide fr
36	35	100.0	95	16	AAW72064	Partial peptide fr
37	35	100.0	95	16	AAW72065	Partial peptide fr
38	35	100.0	95	16	AAW72058	Partial peptide fr
39	35	100.0	95	16	AAW72059	Partial peptide fr
40	35	100.0	95	16	AAW72060	Partial peptide fr
41	35	100.0	100	13	AAW25324	Human 5' EST relat
42	35	100.0	100	21	AAV64671	Light chain VK10.1
43	35	100.0	101	13	AAW22577	IgG antibody 2.1.1
44	35	100.0	102	21	AAV34316	Humanized 369 11gh
45	35	100.0	103	15	AAW49333	Light chain region

ALIGNMENTS

RESULT 1

AAV41874

ID AAV41874 standard; Peptide: 16 AA.

AAV41874:

09-DEC-1999 (first entry)

Rheumatoid arthritis diagnostic protein isoform peptide #25.

DE Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;

KW rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;

KW Rheumatoid arthritis diagnostic protein isoform; screening;

KW expression reference protein isoform; prognosis.

XX Homo sapiens.

OS WO9947925-A2.

XX PD 23-SEP-1999.

XX PF 15-MAR-1999; 99WO-GB00763.

XX PR 13-MAR-1998; 98GB-0005477.

XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX PI Parekh RB, Patel TP, Townsend RR;

XX DR WPI: 1999-571871/48.

XX PT Diagnosis of human rheumatoid arthritis by two-dimensional

XX electrophoresis -

PS Disclosure; Page 18; 157pp; English.

XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a
CC two-dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADf) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADfs can be used in gene therapy
CC protocols. AA141844 to AA142100 represent RPI peptides, AA142101 to
CC AA142103 represent expression reference protein isoform peptides and
CC AA225066 to AA225068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention.

XX
XX Sequence 16 AA:

Query Match 100.0%; Score 35; DB 20; Length 16;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
| | | | |
Db 8 tdfcltl 14

RESULT 2
AAU25389
ID AAU25389 standard; Peptide; 16 AA.

XX
AC AAU25389;

XX
BT 18-DEC-2001 (first entry)

XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #618.

XX
KW Schizophrenia-associated protein isoform; SPI: SPI-206; SPI-238; SPI-240;
KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.

XX
OS Homo sapiens.

XX
PN WO200162785-A2.

XX
PD 30-AUG-2001.

XX
PE 23-FEB-2001; 2001MO-GB00792.

XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.

XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX
PI Herath HMAc, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
XX WPI: 2001-570624/64.

XX
XX New schizophrenia associated protein isoforms and encoding nucleic acid
XX molecules, useful for treatment, diagnosis and prognosis of
XX schizophrenia and screening for potential drugs for treatment and new
XX drug targets -

PS Disclosure; Page 41; 148pp; English.

XX
XX The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs.

XX
XX Sequence 16 AA:

Query Match 100.0%; Score 35; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
| | | | |
Db 8 tdfcltl 14

RESULT 3
AAU15733
ID AAU15733 standard; Peptide; 16 AA.

XX
AC AAU15733;

XX
BT 24-OCT-2001 (first entry)

XX
DE Schizophrenia-associated isoform peptide #618.

XX
KW Schizophrenia: neuroleptic; diagnostic; neuropsychiatric disorder;
KW neurological disorder; neuropathy.

XX
OS Homo sapiens.

XX
PN WO200163293-A2.

XX
PD 30-AUG-2001.

XX
PE 23-FEB-2001; 2001MO-GB00783.

XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.

XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX
PI Herath HMAc, Parekh RB, Rohlf C;
XX WPI: 2001-502868/55.

XX
DR WPI: 2001-502868/55.

XX
PT Diagnosing and monitoring Schizophrenia by detecting the presence of
PT Schizophrenia Associated Features and Schizophrenia Associated Protein
PT Isoforms in samples of cerebrospinal fluid -

XX
PS Claim 6; Page 41; 160pp; English.

XX
XX The invention relates to methods and compositions for screening,
CC diagnosis and prognosis of Schizophrenia. The method involves detecting
CC the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH
CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC studying the effectiveness of treatments and for identifying potential
CC therapeutic agents. The method is used for (1) screening or diagnosis of
CC SCH and the relative abundance of at least 1 chosen feature correlates
CC with the presence or absence of SCH; and (2) monitoring the effect of
CC therapy administered to a subject with SCH and the relative abundance of
CC at least 1 chosen feature which correlates with the severity of SCH.
CC The expression and activity of the SFS, SPIs and related molecules
CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the
CC progress of the disorder and the effectiveness of treatment and as

CC targets to identify and produce potential therapeutic agents for the
 CC treatment of SCH. The paucity of detectable neuralgic defects
 CC distinguishes neuropsychiatric disorders such as SCH from neurological
 CC disorders, where manifestations of anatomical and biochemical changes
 CC have been identified in many cases. Consequently the identification and
 CC characterisation of cellular and/or molecular causative defects and
 CC neuropathies are necessary for improved treatment of neuropsychiatric
 CC disorders. AA015114-AA015762 represent the amino acid sequences of
 CC schizophrenia-associated isoforms used in the method of the invention.
 XX
 SQ Sequence 16 AA;

Query Match 100.0%; Score 35; DB 22; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TDFTLTI 7
 |||||
 Db 8 tdfitli 14

RESULT 4

AA87043
 ID AA87043 standard; Peptide; 32 AA.
 XX
 AC AA87043;
 XX
 DT 25-JUN-1996 (first entry)
 XX
 DE Human group I light chain framework 3.
 XX
 KM Humanised antibody; interleukin-5; IL-5; recombinant antibody;
 KM antibody engineering; monoclonal antibody; MAb; 39D10; CDR;
 KW complementarity determining region; light chain; framework;
 KW eosinophilia; allergy; asthma.
 XX
 OS Homo sapiens.
 XX
 PN WO9535375-A1.
 XX
 PD 28-DEC-1995.
 XX
 PF 16-JUN-1995; 95WO-GB01411.
 XX
 PR 17-JUN-1994; 94GB-0012230.
 XX
 PA (CLLT) CELLTECH THERAPEUTICS LTD.
 XX
 PI Athwal DS, Bodmer MW, Emtage JS;
 XX
 DR WPI; 1996-058412/06.
 XX
 PT Anti-human IL-5 recombinant antibody - useful for preventing or
 PT reducing eosinophilia and for treating certain allergic diseases,
 PT esp. asthma
 XX
 PS Example 3; Fig 3; 69pp; English.
 XX
 CC Framework regions (AA87041-44) of human group I (gpl) germ line
 CC antibody light chain showed homology to corresponding regions
 CC (AA87045-48, respectively) of the rat anti-human interleukin-5
 CC monoclonal antibody 39D10 light chain (see AA87040). This homology
 CC was utilised in the prodn. of a humanised 39D10 VL (AA87057) in
 CC which rat 39D10 VL complementarity determining regions were grafted
 CC into the human gpl framework.
 CC
 XX
 SQ Sequence 32 AA;

Query Match 100.0%; Score 35; DB 17; Length 32;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
 |||||
 Db 13 tdfitli 19

RESULT 5

AA52745
 ID AA52745 standard; Peptide; 32 AA.
 XX
 AC AA52745;
 XX
 DT 26-JAN-2000 (first entry)
 XX
 DE Humanised ATR-5 L chain V region FR3 for "a".
 XX
 KM Human tissue factor; TF; humanised; antibody; mouse monoclonal antibody;
 KM ATR-2; ATR-3; ATR-4; ATR-5; ATR-7; ATR-8; thrombotic disease; DIC;
 KW disseminated intravascular coagulation; immunogenicity; chimeric.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO951743-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 02-APR-1999; 99WO-JP01768.
 XX
 PR 03-APR-1998; 98JP-0091850.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Sato K, Adachi H, Yabuta N;
 XX
 DR WPI; 1999-620204/53.
 XX
 PT Humanised antibody recognizing human tissue factor, used for treatment
 PT of disseminated intravascular coagulation -
 XX
 PS Claim 17; Page 270; 291pp; Japanese.
 XX
 CC The present invention describes chimeric antibody (Ab) heavy (H) chains
 CC containing the variable region of the H chain of a mouse monoclonal Ab
 CC recognising human tissue factor (HTF) and the constant region of the H
 CC chain of a human Ab. The variable region is one of six specified
 CC sequences (which are the H chain variable regions from mouse monoclonal
 CC Ab's ATR-2,3,4,5,7 or 8). Also described are chimeric Ab light (L)
 CC chains containing the variable region of the L chain of a mouse
 CC monoclonal Ab recognising human tissue factor (HTF) and the constant
 CC region of the L chain of a human Ab, the variable region being one of six
 CC specified sequences (which are the L chain variable regions from mouse
 CC monoclonal Ab's ATR-2,3,4,5,7 or 8). The chimeric Ab's can be used for
 CC the treatment and prevention of thrombotic disease, especially of
 CC disseminated intravascular coagulation (DIC). The humanised antibody has
 CC the high htf binding activity of the mouse monoclonal antibody but
 CC greatly reduced immunogenicity. AA233001 to AA233091 and Y527007 to
 CC AA52767 represent sequences used in the exemplification of the present
 CC invention.
 CC
 XX
 SQ Sequence 32 AA;

Query Match 100.0%; Score 35; DB 20; Length 32;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TDFTLTI 7
 |||||
 Db 13 tdfitli 19

RESULT 6

```

ID AAB98286
AC AAB98286 standard; Peptide: 32 AA.
XX
XX AAB98286;
XX
XX 20-AUG-2001 (first entry)
XX
XX
XX Anti-A33 antigen immunoglobulin VL FR3 SEQ ID NO:92.
XX
XX Human; rabbit; humanised; A33 antigen; anti-A33 antigen antibody;
XX immunoglobulin; complementarily determining region; CDR; cancer;
XX
XX
XX
XX Homo sapiens.
XX
XX WO200130393-A2.
XX
XX 03-MAY-2001.
XX
XX
XX 20-OCT-2000; 2000WO-US29289.
XX
XX 22-OCT-1999; 98US-0425638.
XX 04-APR-2000; 2000US-0543004.
XX
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX (SLOK) SLOAN KETTERING INST CANCER RES.
XX (SCRI) SCRIPPS RES INST.
XX
XX
XX Barbas CF, Rader C, Rittler G, Welt S, Old LJ;
XX
XX WPI; 2001-328613/34.
XX
XX Treating cancers, particularly of stomach and colon, that express A33
XX antigen by administering conjugate of anticancer agent with specific
XX immunoglobulin product -
XX
XX Claim 16; Page 40; 85pp; English.
XX
XX The present invention describes a method for treating cancers that
XX express the A33 antigen. The method comprises administering an
XX anticancer agent (I) conjugated to an immunoglobulin product (II) that
XX binds specifically to A33 and contains one or more of 13 specified
XX complementarily determining regions (CDRs), given in AAB98262 to
XX AAB98274. (I) has cytostatic activity. The method can be used for
XX treating colon and stomach cancers. (II), or the nucleic acid encoding
XX it, can be used directly, in unconjugated form, for immunotherapy of
XX cancer, and, when labeled, for detection or diagnosis of diseases
XX associated with A33 expression. AAH22218 to AAH22254 and AAB98230 to
XX AAB98231 represent sequences used in the exemplification of the
XX present invention.
XX
XX Sequence 32 AA;
XX

```

```

Query Match      100.0%; Score 35; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 2,1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TDEFTTIT 7
          |||||
DB      13 tdficli 19

RESULT 7
AAB98288
ID AAB98288 standard; Peptide: 32 AA.
XX
AC AAB98288;
XX
XX 20-AUG-2001 (first entry)
DT
XX
DE Ant1-A33 antigen immunoglobulin VL FR3 SEQ ID NO:94.
XX

```

OS Homo sapiens.
PN WO200130393-A2.
XX
XX
PD 03-MAY-2001.
XX
XX
PF 20-OCT-2000; 2000WO-US29289.
XX
PR 22-OCT-1999; 99US-0425638.
XX 04-APR-2000; 2000US-0543004.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PA (SCRI) SCRIPPS RES INST.
X1 Barbas CF, Rader C, Ritter G, Welt S, Old LJ:
XX
XX WPI; 2001-328613/34.
DR
XX
PT Treating cancers, particularly of stomach and colon, that express A33
PT antigen by administering conjugate of anticancer agent with specific
PT immunoglobulin product -
XX
XX
PS Claim 16; Page 40; 85pp; English.

The present invention describes a method for treating cancers that express the A33 antigen. The method comprises administering an anticancer agent (I) conjugated to an immunoglobulin product (II) that binds specifically to A33 and contains one or more of 13 specified complementarily determining regions (CDRs), given in AAB98262 to AAB98274. (I) has cytostatic activity. The method can be used for treating colon and stomach cancers. (II), or the nucleic acid encoding it, can be used directly, in unconjugated form, for immunotherapy of cancer, and, when labeled, for detection or diagnosis of diseases associated with A33 expression. AAH22218 to AAH22254 and AAP98230 to AAP98321 represent sequences used in the exemplification of the present invention.

Sequence 32 AA;

```

Query Match      100.0%; Score 35; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. NO. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TDFPLTIT 7
          |||||
Db       13 tdfllti 19

RESULT      8
AAB97666
1D AAB97666 standard; Peptide; 32 AA.
XX
XX
XX AAB97666;
XX
XX
XX 08-AUG-2001 (first entry)
XX
XX
XX A33 antigen binding immunoglobulin product VLFR3 peptide SEQ ID NO:92.
XX
XX Chimeric antibody; humanised; humanisation; mammalian antibody; antigen;
XX immunoreact; anti A33 antigen antibody; immunoglobulin.
XX
XX Homo sapiens.
XX
XX WO200131065-A1.
XX
XX
XX 03-MAY-2001.
XX
XX

```


QY	1	TDFLTl	7
Db	13	tdfllt	19
RESULT	9		
AAB97668			
ID	AAB97668	standard; Peptide; 32 AA.	
XX			
AC	AAB97668;		
DT	08-AUG-2001	(first entry)	
XX			
DE	A33 antigen binding immunoglobulin product VLEF3 peptide SEQ ID NO:94.		
XX			
KM	Chimeric antibody; humanised; humanisation; mammalian antibody; antigen; immunoreact; anti A33 antigen antibody; immunoglobulin.		
OS	Homo sapiens.		
XX			
PN	WO200131065-A1.		
PD	03-MAY-2001.	1	
Pf	20-OCT-2000; 2000MO-US29026.		
PR	22-OCT-1999; 99US-0425638.		
PA	04-APR-2000; 2000US-0543004.		
PI	(SCRI) SCRIPPS RES INST.		
DR	Barbas CF, Rader C;		
WPI	2001-328657/34.		
Preparing humanized rabbit antibodies that specifically immunoreact with a particular antigen using display technology for expressing			

Query Match	100.0%;	Score 35;	DB 22;	Length 32;
Best Local Similarity	100.0%;	Pred. No. 2.1;		
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

RESULT 10

AC AAW62805;

DT 23-SEP-1998 (first entry)

Amino acid sequence of a human antibody fragment.

AA Human; immunoglobulin; Ig; transgenic; non-human mammal;
KW

inactivated endogenous Ig locus; B-cell development; KW

KW human heavy chain Ig locus; J-H; D-H; V-H gene;

κ₁ kappa light chain Ig locus; κ₂ kappa constant region; J-kappa gene; V-kappa;

KW production; antibody.

XX

OS Homo sapiens.

XX

PN W09824893-A2.

XX 11 1111-1000

PD 11-JUN-1998.
XX

XX
DE 03-DEC-1997. 07WD-11523091

PF 03-DEC-1997; 9/WO-0523091.
XY

XX
PR 03-DEC-1996: 96JJS-0759620

PR 03-DEC-1990; 3005
XX

AA
PA (ABGE-) ABGENIX INC.

PA (ABOE) MBOENIA INC
XX

Green L, Jakobovits A, Klapholz S, Kucherlapati R; PI

PI Mendez M;

XX 7

DR WPI; 1998-333314/29.

XX

PT New transgenic non-h

PT immunoglobulin locus

PT used for production

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

PS Disclosure; Page 77;

XX
XX
DATE 07-03-00 000000

CC	AAW62793-822	representative
CC	AAW62793-822	representative

transgenic Xenomice, CC specification descr:

CC specification descriptions

modifications that

CC locus, so that the
CC modified genome also

...genetic pattern ...

germline configuration, the human heavy chain Ig locus comprising a human micro constant region and regulatory and switch sequences, human J-H genes, human D-H genes, and human V-H genes and an inserted human kappa light chain Ig locus in germline configuration, the human kappa light chain Ig locus comprising a human kappa constant region, J-kappa genes, and V-kappa genes, where the number of V-H and V-kappa genes inserted are selected to restore normal B-cell development in the mammal. The transgenic animals have a near complete human Ig locus, including both a human heavy chain locus and a human kappa light chain locus. They can be used for the production of human antibodies when exposed to particular antigens e.g. When exposed to human IL-8, EGFR or TNF- alpha the mice will produce antibodies to IL-8, EGFR or TNF- alpha respectively.

Query Match	100.0%	Score 35;	DB 19;	Length 74;					
Best Local Similarity	100.0%;	Pred. No. 5.1;							
Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;

```

QY      1 TDFLT 7
Db      48 tdfc1 54

```

RESULT 11
AAW80981
ID AAW80981 standard; Protein; 76 AA.

AC AAW80981;

DT 30-MAR-1999 (first entry)

Variable kappa 11gnt region 012 encoded amino acid.

KW Human; epidermal growth factor receptor; tumour; EGF; carcinoma

transforming

OS Homo sapiens.

PN W09850433-Å2.

PD 12-NOV-1998.

PF 05-MAY-1998; 98MO-US09160.

PR 05-MAY-1997; 97US-0851362.

PA (ABGE-) ABGENIX INC.

PI Gallo M, Jakobovits A, Jia X, Yang X;

DR WPI; 1999-034712/03.

PT Humanised antibodies

PT or allergic effects

PS Example 3; Page 105; 143pp; English.

The variable kappa light region 012 encoded amino acid was used in the production of anti-epidermal growth factor receptor (Egfr-1)-antibodies. The antibodies can be administered therapeutically to patients (human or veterinary) to treat solid tumours. Egfr-1 is overexpressed on many human solid tumour types, and the fully human antibodies (i.e. comprising and inhibit both epidermal growth factor (EGF) and transforming growth factor alpha (TGF-alpha) binding to Egfr-1 (known to lead to cellular proliferation and tumour growth). They can prevent tumour cell growth and, in combination with an antineoplastic agent, may eradicate established tumours. The fully human antibodies can minimise the immunogenic and allergic responses intrinsic to previous mouse/rat or mouse/rat-derived antibodies.

XX	Sequence	76 AA;
SQ		

Query Match	100.0%	Score 35;	DB 20;	Length 76;
Best Local Similarity	100.0%	Pred. No. 5.2;		
Matches	7;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	1	TDFTLTI	7
Db	50	tdftlti	56

RESULT	12
AAW62807	
ID	AAW62807 standard; Peptide: 82 AA

AC AAW62807;

DT 23-SEP-1998 (first entry)

Amino acid sequence of a human antibody fragment.

KW Human; immunoglobulin; Ig; transgenic; non-human mammal;

KW human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;

production; antibody.

05 Homo sapiens.

PN W09824893-A2.

PD 11-JUN-1998.

PF 03-DEC-1997; 97WO-US23091.

PR 03-DEC-1996; 96US-0759620.

PA (ABGE-) ABGENIX INC.

PI Green L, Jakobovits A, Klapholz S, Kucherlapati R,

XX

XX

immunoglobulin locus and a near complete human immunoglobulin locus.

XX

XX
XX

AA6227933232 represent fragments of human antibodies produced by transgenic xenomice, created using the method of the invention. The specification describes a transgenic non-human mammal which has genome modifications that comprise an inactivated endogenous immunoglobulin (Ig) locus, so that the mammal does not display normal B-cell development. The modified genome also has an inserted human heavy chain Ig locus in germline configuration, the human heavy chain Ig locus comprising a human constant region and regulatory and switch sequences, human J-H genes, human D-H genes, and human V-H genes and an inserted human kappa light chain Ig locus in germline configuration, the human kappa light chain Ig locus comprising a human kappa constant region, J-kappa genes, and V-kappa genes, where the number of V-H and V-kappa genes inserted are selected to restore normal B-cell development in the mammal. The transgenic animals have a near complete human Ig locus, including both a human heavy chain locus and a human kappa light chain locus. They can be used for the production of human antibodies when exposed to particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha the mice will produce antibodies to IL-8, EGFR or TNF- alpha respectively.

Query Match 100.0%; Score 35; DB 19; Length 82;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 |||||
 DB 56 tdfilti 62

RESULT 13

AAW14491
 ID AAW14491 standard; Protein; 84 AA.

AC AAW14491;

DT 28-JAN-1997 (first entry)

DE Monoclonal antibody D VK.

heavy; light chain; monoclonal antibody; antigen 3; marker; melanoma;

permanent human tumour cell line; tumour-associated antigen; epitope;

gastrointestinal tumour; pancreatic carcinoma; diagnostic; therapeutic;

antigen 11; Vibrio cholera; neuraminidase-resistant; ganglioside GD2.

Synthetic.

EP727436-A1.

21-MAR-1990; 90EP-0105322.

24-MAR-1989; 89DE-3909799.

(BEHM) BEHRINGWERKE AG.

Auerbach B, Bosslet K, Sedlacek H, Seemann G;

WPI: 1996-372836/38.

N-PSDB; AAT63508.

Monoclonal antibody to tumour-associated antigen - useful as

gastrointestinal tumour marker

Disclosure: Page 14; 19pp; German.

AAW14490-91 are the heavy and light chains (respectively) of monoclonal

antibody (Mab) D. Mab D recognises Vibrio cholera

neuraminidase-resistant epitope of ganglioside GD2, from a human melanoma

cell line. Mabs A, B and C (see AAW14484-89) are mentioned in the

specification, but are not part of the claims. Mabs A and B recognise

antigens 3 and 11 resp. of a permanent human tumour cell line. Mab C

also recognises an epitope of a tumour-associated antigen. These antigens

occur at high concns. in the serum of patients with gastrointestinal

tumours, e.g. pancreatic carcinoma, and are thus useful as tumour markers

for diagnostic or therapeutic purposes.

Sequence 84 AA;

Query Match 100.0%; Score 35; DB 17; Length 84;

Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID AAR9878 standard; Protein; 84 AA.

AC AAR9878;

DT 28-JAN-1997 (first entry)

DE Monoclonal antibody D VK.

Monoclonal antibody; Mab; epitope; tumour-associated antigen;

marker; antigen.

Synthetic.

EP727435-A1.

21-AUG-1996.

21-MAR-1990; 90EP-0105322.

24-MAR-1989; 89DE-3909799.

(BEHM) BEHRINGWERKE AG.

Auerbach B, Bosslet K, Sedlacek H, Seemann G;

WPI: 1996-372835/38.

N-PSDB; AAT36666.

Monoclonal antibody to tumour-associated antigen - useful as

gastrointestinal tumour marker

Disclosure: Page 14; 19pp; German.

Mab C (AAT36659-T36660) is a monoclonal antibody that recognises an

epitope of a tumour-associated antigen occurring at high concn. in

the serum of patients with gastrointestinal tumours, e.g. pancreatic

carcinoma, and is thus useful as a tumour marker for diagnostic or

therapeutic purposes.

Mabs A, B and D are mentioned in the specification, but are not

part of the claims.

Mab A (AAT36661-T36662) recognises antigen 3 of permanent human

tumour cell line.

Mab B (AAT36663-T36664) recognises antigen 11 of permanent human

tumour cell line.

Mab D (AAT36665-T36666) recognises a Vibrio cholera neuraminidase-

resistant epitope of ganglioside GD2, from a human melanoma cell

line.

Sequence 84 AA;

Query Match 100.0%; Score 35; DB 17; Length 84;

Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
 |||||

DB 47 tdfilti 53

RESULT 15

AAW62806
 ID AAW62806 standard; Peptide; 86 AA.

AC AAW62806;

DT 23-SEP-1998 (first entry)

DE Amino acid sequence of a human antibody fragment.

Human; immunoglobulin; Ig; transgenic; non-human mammal;

inactivated endogenous Ig locus; B-cell development;

human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;

KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
 KW production; antibody.
 XX
 OS Homo sapiens.
 XX
 PN WO9824893-A2.
 XX
 PD 11-JUN-1998.
 XX
 PF 03-DEC-1997; 97WO-US23091.
 XX
 PR 03-DEC-1996; 96US-0759620.
 XX
 PA (ABGE-) ABGENIX INC.
 XX
 PI Green I, Yakobovits A, Klapholz S, Kucherlapati R;
 PI Mendez M;
 XX
 DR WPI; 1998-333314/29.
 XX
 PT New transgenic non-human mammals - having an inactivated
 PT immunoglobulin locus and a near complete human immunoglobulin locus,
 PT used for production of human antibodies
 XX
 PS Disclosure; Page 78; 128pp; English.
 XX
 CC AAM62793-822 represent fragments of human antibodies produced by
 CC transgenic Xenomice, created using the method of the invention. The
 CC specification describes a transgenic non-human mammal which has genome
 CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
 CC locus, so that the mammal does not display normal B-cell development. The
 CC modified genome also has an inserted human heavy chain Ig locus in
 CC germline configuration, the human heavy chain Ig locus comprising a human
 CC micro constant region and regulatory and switch sequences, human J-H
 CC genes, human D-H genes, and human V-H genes and an inserted human kappa
 CC light chain Ig locus in germline configuration, the human kappa light
 CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
 CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
 CC are selected to restore normal B-cell development in the mammal. The
 CC transgenic animals have a near complete human Ig locus, including both a
 CC human heavy chain locus and a human kappa light chain locus. They can
 CC be used for the production of human antibodies when exposed to
 CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
 CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
 CC respectively.
 CC
 SQ Sequence 86 AA;
 XX
 QY 1 TDTFTLT 7
 DB 48 tdtftlti 54
 XX

Search completed: July 15, 2002, 12:57:57
 Job time: 413 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:59:18 ; Search time 75.67 Seconds
(without alignments)
2.260 Million cell updates/sec

Title: US-09-712-819A-5

Perfect score: 1 TDFTLTI 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUTS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	25	5	PCT-US91-02942-42
2	35	100.0	32	2	US-08-470-139-13
3	35	100.0	32	4	US-08-569-147-40
4	35	100.0	32	4	US-09-347-061-13
5	35	100.0	32	4	US-09-425-638A-92
6	35	100.0	32	4	US-09-425-638A-94
7	35	100.0	32	4	US-09-543-004-92
8	35	100.0	32	4	US-09-543-004-94
9	35	100.0	50	5	PCT-US91-02942-9
10	35	100.0	52	1	US-08-162-102C-43
11	35	100.0	53	1	US-08-162-102C-44
12	35	100.0	64	2	US-08-765-179B-10
13	35	100.0	64	2	US-08-765-179B-14
14	35	100.0	70	3	US-08-554-840-9
15	35	100.0	76	4	US-08-851-362D-21
16	35	100.0	79	3	US-08-554-840-14
17	35	100.0	80	3	US-08-554-840-12
18	35	100.0	80	3	US-08-554-840-13
19	35	100.0	80	3	US-08-554-840-15
20	35	100.0	81	3	US-08-554-840-11
21	35	100.0	91	2	US-08-273-146-49
22	35	100.0	93	1	US-08-276-852-111
23	35	100.0	93	1	US-08-899-575-111
24	35	100.0	93	1	US-08-899-575-111
25	35	100.0	93	3	US-08-783-853A-35
26	35	100.0	93	5	PCT-US95-08743-111
27	35	100.0	96	3	US-08-466-368-6

28	35	100.0	103	4	US-09-240-274-42	Sequence 42, Appl
29	35	100.0	104	1	US-08-276-852-92	Sequence 92, Appl
30	35	100.0	104	1	US-08-276-852-94	Sequence 94, Appl
31	35	100.0	104	1	US-08-276-852-100	Sequence 100, App
32	35	100.0	104	1	US-08-276-852-106	Sequence 106, App
33	35	100.0	104	1	US-08-899-575-92	Sequence 92, Appl
34	35	100.0	104	1	US-08-899-575-94	Sequence 94, Appl
35	35	100.0	104	1	US-08-899-575-100	Sequence 100, App
36	35	100.0	104	1	US-08-899-575-106	Sequence 106, App
37	35	100.0	104	1	US-08-899-575-92	Sequence 92, Appl
38	35	100.0	104	1	US-08-899-575-94	Sequence 94, Appl
39	35	100.0	104	1	US-08-899-575-100	Sequence 100, App
40	35	100.0	104	1	US-08-899-575-106	Sequence 106, App
41	35	100.0	104	5	PCT-US95-08743-92	Sequence 92, Appl
42	35	100.0	104	5	PCT-US95-08743-94	Sequence 94, Appl
43	35	100.0	104	5	PCT-US95-08743-100	Sequence 100, App
44	35	100.0	104	5	PCT-US95-08743-106	Sequence 106, App
45	35	100.0	105	1	US-08-276-852-89	Sequence 89, Appl

ALIGNMENTS

RESULT 1
PCT-US91-02942-42
; Sequence 42, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: AHWAL, DILEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02942
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009549.8
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAM L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 1011.0586600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US91-02942-42

Query Match 100.0%; Score 35; DB 5; Length 25;
Best Local Similarity 100.0%; Pred No. 0.68; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
Oy 1 TDFTLTI 7
|||||||

Db 6 TDFLT1 12

RESULT 2

US-08-470-139-13

; Sequence 13, Application US/08470139

; Patent No. 5998586

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: Interleukin-5 specific recombinant antibodies

; NUMBER OF SEQUENCES: 28

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/470,139

; FILING DATE: 06 JUNE-1995

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: TRUJILLO, DOREEN YATKO

; REGISTRATION NUMBER: 35,719

; REFERENCE/DOCKET NUMBER: CARP-0044

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 32 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-470-139-13

Query Match 100.0%; Score 35; DB 2; Length 32;

Best Local Similarity 100.0%; Pred. No. 0.88; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFLT1 7

Db 13 TDFLT1 19

RESULT 3

US-08-569-147-40

; Sequence 40, Application US/08569147

; Patent No. 6180377

; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: HUMANISED ANTIBODIES

; NUMBER OF SEQUENCES: 95

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

; ADDRESSEE: No. 61803771s, LLP

; STREET: One Liberty Place - 46th Floor

; CITY: Philadelphia

; STATE: PA

; COUNTRY: U.S.A.

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.26 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/569,147

; FILING DATE: 25-March-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Trujillo, Doreen Yatko

; REGISTRATION NUMBER: 35,719

; REFERENCE/DOCKET NUMBER: CARP-0047

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439

; INFORMATION FOR SEQ ID NO: 40:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 32 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-569-147-40

Query Match 100.0%; Score 35; DB 4; Length 32;

Best Local Similarity 100.0%; Pred. No. 0.88; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFLT1 7

Db 13 TDFLT1 19

RESULT 4

US-09-347-061-13

; Sequence 13, Application US/09347061

; Patent No. 6316227

; GENERAL INFORMATION:

; APPLICANT: Bodmer, Mark

; APPLICANT: Athwal, Diljeet Singh

; TITLE OF INVENTION: Emtage, John Spencer

; FILE REFERENCE: Interleukin-5 Specific Recombinant Antibodies

; CURRENT APPLICATION NUMBER: US/09/347,061

; CURRENT FILING DATE: 1999-07-02

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 13

; LENGTH: 32

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc.feature

; OTHER INFORMATION: Consensus

US-09-347-061-13

Query Match 100.0%; Score 35; DB 4; Length 32;

Best Local Similarity 100.0%; Pred. No. 0.88; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFLT1 7

Db 13 TDFLT1 19

RESULT 5

US-09-425-638A-92

; Sequence 92, Application US/09425638A

; Patent No. 6342587

; GENERAL INFORMATION:

; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt an

; APPLICANT: Lloyd J. Old A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THE

; TITLE OF INVENTION: LUD 5630

; FILE REFERENCE: US/09/425,638A

; CURRENT APPLICATION NUMBER: US/09-10-22

; CURRENT FILING DATE: 1999-10-22

; NUMBER OF SEQ ID NOS: 129

; SEQ ID NO 92

; LENGTH: 32

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

US-09-425-638A-92

Query Match 100.0%; Score 35; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||||
DB 13 TDFTLTI 19

RESULT 6
US-09-425-638A-94
; Sequence 94, Application US/09425638A
; Patent No. 6342587
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630.1
; CURRENT APPLICATION NUMBER: US/09/425,638A
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 94
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-425-638A-94

Query Match 100.0%; Score 35; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||||
DB 13 TDFTLTI 19

RESULT 7
US-09-543-004-92
; Sequence 92, Application US/09543004
; Patent No. 6346249
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630.1
; CURRENT APPLICATION NUMBER: US/09/543,004
; CURRENT FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: 09/425,638
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 92
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-543-004-92

Query Match 100.0%; Score 35; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||||
DB 13 TDFTLTI 19

RESULT 8
US-09-543-004-94
; Sequence 94, Application US/09543004

Patent No. 6346249
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630.1
; CURRENT APPLICATION NUMBER: US/09/543,004
; CURRENT FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: 09/425,638
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 94
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-543-004-94

Query Match 100.0%; Score 35; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
|||||||
DB 13 TDFTLTI 19

RESULT 9
PCT-US91-02942-9
; Sequence 9, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: ARTHAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Ave. NW Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02942
FILING DATE: 19910429
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9009549.8
FILING DATE: 27-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAM L
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1011.0586600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US91-02942-9

Query Match 100.0%; Score 35; DB 5; Length 50;

Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
111111
Db 24 TDFTLTI 30

RESULT 10
US-08-162-102C-43

; Sequence 43, Application US/08162102C
; Patent No. 5762905

; GENERAL INFORMATION:

; APPLICANT: Burton, Dennis R.

; APPLICANT: Barbas, III, Carlos F.

; APPLICANT: Chanock, Robert M.

; APPLICANT: Murphy, Brian R.

; APPLICANT: Crowe, Jr., James E.

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES

; TITLE OF INVENTION: TO RESPIRATORY SYNCYTIAL VIRUS

; NUMBER OF SEQUENCES: 49

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: California

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/162,102C

; FILING DATE: 10-DEC-1993

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Hallie, Ph.D., Lisa A.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07300/007001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 52 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-162-102C-43

QY 1 TDFTLTI 7
111111
Db 13 TDFTLTI 19

RESULT 11
US-08-162-102C-44

; Sequence 44, Application US/08162102C
; Patent No. 5762905

; GENERAL INFORMATION:

; APPLICANT: Burton, Dennis R.

; APPLICANT: Barbas, III, Carlos F.

; APPLICANT: Chanock, Robert M.

; APPLICANT: Murphy, Brian R.

; APPLICANT: Crowe, Jr., James E.

; TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
; TITLE OF INVENTION: TO RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: California

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/162,102C

; FILING DATE: 10-DEC-1993

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Hallie, Ph.D., Lisa A.

; REGISTRATION NUMBER: 38,347

; REFERENCE/DOCKET NUMBER: 07300/007001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 44:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 53 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-162-102C-44

Query Match 100.0%; Score 35; DB 1; Length 53;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTLTI 7
111111
Db 13 TDFTLTI 19

RESULT 12
US-08-765-179B-10
; Sequence 10, Application US/08765179B
; Patent No. 3834027

; GENERAL INFORMATION:

; APPLICANT: STEIPE, Boris

; APPLICANT: STEINBACHER, Stefan

; TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY

; TITLE OF INVENTION: OF ANTIBODIES

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP

; STREET: 655 Fifteenth Street N.W. Suite 330

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3701

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/765,179B

; FILING DATE: 14-JAN-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/EP95/02626

;; FILING DATE: 06-JUL-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DE P 44 25 115.7
;; FILING DATE: 15-JUL-1994
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 64 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-765-179B-10

Query Match 100.0%; Score 35; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
DB 38 TDFTLTI 44

RESULT 13
US-08-765-179B-14
; Sequence 14, Application US/08765179B
; Patent No. 5854027
; GENERAL INFORMATION:
; APPLICANT: STEIPE, Boris
; APPLICANT: STEINBACHER, Stefan
; TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY
; TITLE OF INVENTION: OF ANTIBODIES
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,179B
; FILING DATE: 14-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/02626
; FILING DATE: 06-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 25 115.7
; FILING DATE: 15-JUL-1994
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-765-179B-14

Query Match 100.0%; Score 35; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
DB 38 TDFTLTI 44

RESULT 14
US-08-554-840-9
; Sequence 9, Application US/08554840
; Patent No. 6001358
; GENERAL INFORMATION:
; APPLICANT: BLACK, Amelia
; APPLICANT: HANNA, Nabil
; APPLICANT: PADLAN, Eduardo A.
; APPLICANT: NEMAN, Roland A.
; TITLE OF INVENTION: HUMANIZED ANTIBODIES TO HUMAN gp39,
; TITLE OF INVENTION: COMPOSITIONS CONTAINING AND THERAPEUTIC USE THEREOF
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 2213-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,840
; FILING DATE: 07-NOV-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Teskin, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 012712-127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 70 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-554-840-9

Query Match 100.0%; Score 35; DB 3; Length 70;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
|||||
DB 51 TDFTLTI 57

RESULT 15
US-08-851-362D-21
; Sequence 21, Application US/08851362D
; Patent No. 6235883
; GENERAL INFORMATION:
; APPLICANT: Jakobovits, Aya
; APPLICANT: Yang, Xiao-Dong
; APPLICANT: Gallo, Michael
; APPLICANT: Jia, Xiao-Chi
; TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal
; TITLE OF INVENTION: Growth Factor Receptor
; FILE REFERENCE: Cell 4.20
; CURRENT APPLICATION NUMBER: US/08/851,362D
; CURRENT FILING DATE: 1997-05-05
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 76

; TYPE: PRT
; ORGANISM: human
US-08-851-362D-21

Query Match 100.0%; Score 35; DB 4; Length 76;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TDFTLTI 7
|||||
DB 50 TDFTLTI 56

Search completed: July 15, 2002, 12:59:18
Job time: 389 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:27:08 ; Search time 49.13 Seconds
(Without alignments)
13.691 Million cell updates/sec

Title: US-09-712-819a-5

Perfect score: 35

Sequence: 1 TDFLRLI 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 455

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	37.1	6	2	PD0028
2	12	34.3	5	2	A32516
3	12	34.3	5	2	PT0644
4	12	34.3	7	2	B39127
5	12	34.3	7	2	PT0665
6	11	31.4	4	2	I40697
7	11	31.4	7	2	E30608
8	10	28.6	6	2	PT0650
9	10	28.6	7	2	S09066
10	9	25.7	3	3	T13892
11	9	25.7	5	2	E42364
12	9	25.7	5	2	E60274
13	9	25.7	5	2	S68326
14	9	25.7	5	2	S69237
15	9	25.7	6	2	A60986
16	9	25.7	6	2	B44510
17	9	25.7	6	2	A43766
18	9	25.7	6	2	T17263
19	9	25.7	6	2	S29881
20	9	25.7	6	2	I65546
21	9	25.7	6	2	PT0587
22	9	25.7	7	2	S25266
23	9	25.7	7	2	A25269
24	9	25.7	7	2	A30812
25	9	25.7	7	2	PT0611
26	8	22.9	4	2	PT0697
27	8	22.9	5	2	I39964
28	8	22.9	5	2	I39966
29	8	22.9	5	2	I39965

30	8	22.9	5	2	A44692	fullcin - giant Af
31	8	22.9	5	2	PT0729	T-cell receptor be
32	8	22.9	5	2	PT0590	T-cell receptor be
33	8	22.9	5	2	G44817	27.5 kda structura
34	8	22.9	5	2	I44817	27.5k structural p
35	8	22.9	5	2	E44817	27.5k structural p
36	8	22.9	5	2	C44817	28.5k structural p
37	8	22.9	5	2	A44817	28k structural pro
38	8	22.9	6	2	A19780	transferrin - bovl
39	8	22.9	6	2	A46474	EC epsilon RIIB -
40	8	22.9	6	2	PT0637	T-cell receptor be
41	8	22.9	6	2	PT0641	T-cell receptor be
42	8	22.9	7	2	E61491	seed protein ws-5
43	8	22.9	7	2	PS0254	18k protein 5507 -
44	8	22.9	7	2	PT0642	T-cell receptor be
45	8	22.9	7	2	PT0689	T-cell receptor be

ALIGNMENTS

RESULT 1
PD0028
Pev-Kinln 2 - penaeid shrimp (Penaeus vannamei) (fragment)
C:Species: Penaeus vannamei
C:Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000
C:Accession: PD0028
R:Nieto, J.; Veelaert, D.; Derua, R.; Weelkens, E.; Cersiaens, A.; Coast, G.; Devree
Biochem. Biophys. Res. Commun. 248, 406-411, 1998
A:Title: Identification of one tachykinin- and two kinin-related peptides in the brai
A:Reference number: PD0027; MID:98342103
A:Accession: PD0028
A:Molecule type: protein
A:Residues: 1-6 <NIE>
C:Comment: This peptide belongs to myotropic neuropeptides.

Query Match 37.1%; Score 13; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFT 4
Db 1 DFS 3

RESULT 2
A32516
cholecystokinin-5 - dog
N:Alternate names: CCK-5
C:Species: Canis lupus familiaris (dog)
C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000
C:Accession: A32516
R:Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J
Am. J. Physiol. 252, G272-G275, 1987
A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and in
A:Reference number: A32516; MID:87153871
A:Accession: A32516
A:Molecule type: protein
A:Residues: 1-5 <SHI>
C:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecy
C:Superfamily: gastrin
C:Keywords: amidated carboxyl end; neuropeptide
F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 34.3%; Score 12; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DFT 3
Db 4 DFT 5

RESULT 3
PT0664
T-cell receptor beta chain V-D-J region (111-16) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0644
R:Feeney, A.J.
J:Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601
A:Accession: PT0644
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEBS>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 34.3%; Score 12; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 TDFT 4
: 11
Db 2 SSFT 5

RESULT 4
B39127
phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999
C:Accession: B39127
R:Hardesty, C.; Ferran, C.; DiRienzo, J.M.
J: Bacteriol. 173, 449-456, 1991
A:Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of sc
rin.
A:Reference number: A39127; MUID:91100329
A:Accession: B39127
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-7 <HAB>
A:Cross-references: GB:M38416; NID:9155142; PIDD:AAA98418.1; PID:9155144
C:Keywords: phosphotransferase

Query Match 34.3%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
: 11
Db 2 DF 3

RESULT 5
PT0665
T-cell receptor beta chain V-D-J region (121-38M) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0665
R:Feeney, A.J.
J:Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601
A:Accession: PT0665
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-7 <FEBS>
A:Experimental source: day 4 postnatal thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 34.3%; Score 12; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
: 11
Db 6 DF 7

RESULT 6
I40697
biotin A - Citrobacter freundii (fragment)
C:Species: Citrobacter freundii
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 12-Aug-1996
C:Accession: I40697
R:Shuan, D.; Campbell, A.
Gene 67, 203-211, 1988
A:Title: transcriptional regulation and gene arrangement of Escherichia coli, Citrobac
A:Reference number: I40697; MUID:89006280
A:Accession: I40697
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:M21922; NID:9144434

Query Match 31.4%; Score 11; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TD 2
: 11
Db 2 TD 3

RESULT 7
E30608
Ig kappa chain V-III region (Gag) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996
C:Accession: E30608
R:Gonfi, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; S
J: Immunol. 142, 3158-3163, 1989
A:Title: Structural and idiotypic characterization of the L chains of human IgM autoe
A:Reference number: A30601; MUID:89215279
A:Accession: E30608
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <GON>
C:Keywords: heterotetramer; immunoglobulin

Query Match 31.4%; Score 11; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DFTLT 6
: 11
Db 1 ETLVT 5

RESULT 8
PT0650
T-cell receptor beta chain V-D-J region (121-38F) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0650
R:Feeney, A.J.
J:Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A:Reference number: PT0509; MUID:91277601

A:Accession: P70650
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-6 <FEED>
A:Experimental source: day 4 postnatal thymus, strain BAUB/C
C:Keywords: T-cell receptor

Query Match 28.6%; Score 10; DB 2; Length 6;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 DFT 4
: 1 1
Db 3 SDAT 6

RESULT 9
S09066
globulin IV alpha subunit delta-1 chain, seed - cucurbit (fragments)
N:Alternate names: 115 globulin alpha subunit delta-1 chain
C:Species: Cucurbita sp. (cucurbit)
C>Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 08-Nov-1996
C:Accession: S09066
R:Ohmura, M.; Hara, I.; Matsubara, H.
Plant Cell Physiol. 21, 157-167, 1980
A:Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic and
A:Reference number: S09066
A:Molecule type: protein
A:Accession: S09066
A:Residues: 1-6;7 <OHM>

Query Match 28.6%; Score 10; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 DFT 5
: 1 1
Db 3 DFTI 6

RESULT 10
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion (frag
C:Species: mitochondrion lampetra fluviatilis (river lamprey)
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: T13892
R:Delarbre, C.; Barriol, V.; Tillier, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A:Title: The main features of the craniate mitochondrial DNA between the NDI and the COI
A:Reference number: T13892
A:Accession: T13892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3
A:Cross-references: EMBL:Y09528; NID:g2340016; PIDN:CAA70721.1; PID:g4379123
C:Genetics:
A:Genome: mitochondrion
A:Note: COI
C:Keywords: mitochondrion; oxidoreductase

Query Match 25.7%; Score 9; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TL 5
: 1 1
Db 2 TL 3

RESULT 11

E42364
flagellar protein flir - Salmonella typhimurium (fragment)
C:Species: Salmonella typhimurium
C>Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C:Accession: E42364
R:Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A:Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and
A:Reference number: A42364; MUID:91258342
A:Accession: E42364
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5 <VOG>
A:Cross-references: GB:M62408

Query Match 25.7%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TL 5
: 1 1
Db 3 TL 4

RESULT 12
E60274
major protein antigen MP63 - Mycobacterium tuberculosis (fragment)
C:Species: Mycobacterium tuberculosis
C>Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C:Accession: E60274
R:Nagal, S.; Wilker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A:Title: Isolation and partial characterization of major protein antigens in the cult
A:Reference number: A60274; MUID:91099989
A:Accession: E60274
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <NAG>

Query Match 25.7%; Score 9; DB 2; Length 5;
Best Local Similarity 25.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 FTLT 6
: 1 1
Db 2 YFTI 5

RESULT 13
S68326
blood cell protein B - Ascidia ceratodes (fragment)
N:Alternate names: Abcp-B
C:Species: Ascidia ceratodes
C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 11-Jun-1999
C:Accession: S68326
R:Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A:Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from
A:Reference number: S68325; MUID:96132650
A:Accession: S68326
A:Molecule type: protein
A:Residues: 1-5 <RAY>
F:2/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F:4/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 25.7%; Score 9; DB 2; Length 5;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3

Db 1 DY 2

RESULT 14

S69237
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)
C:Species: Staphylothermus marinus
C:Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 17-Mar-1999
C:Accession: S69237
R:Peters, J.; Nitsch, M.; Kuehlmergen, B.; Goldik, R.; Lupas, A.; Kellermann, J.; Engelh
J. Mol. Biol. 245, 385-401, 1995
A:Title: Tetrabrachion: a filamentous archaeobacterial surface protein assembly of unusua
A:Reference number: S69237; MUID:95139068
A:Accession: S69237
A:Molecule type: protein
A:Residues: 1-5 <P>
A:Experimental source: strain FL, DSM 3639
C:Keywords: cell wall; glycoprotein; heat-stable protein

Query Match

25.7%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TL 5
11
Db 2 TL 3

RESULT 15

A60986
N-formyl oligopeptide - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 31-Dec-1993
C:Accession: A60986
R:Broom, M.F.; Mellor, D.M.; Chadwick, V.S.
Experientia 45, 1097-1099, 1989
A:Title: Purification and amino acid sequencing of naturally occurring N-formyl-methionyl
A:Reference number: A60986; MUID:90092408
A:Accession: A60986
A:Molecule type: protein
A:Residues: 1-6 <BRO>
C:Comment: This hexapeptide was the longest of several N-formyl oligopeptides reported.
F:1/Modified site: N-formylmethionine #status experimental

Query Match

25.7%; Score 9; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 FTL 5
11
Db 3 FTL 5

Search completed: July 15, 2002, 13:27:09
Job time: 439 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:38:10 ; Search time 24.49 Seconds
(without alignments)
11.067 Million cell updates/sec

Title: US-09-712-819A-5

Perfect score: 35
Sequence: 1 TDFLTLT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	40.0	7	1	ALL7_CYPDPO
2	12	34.3	5	1	RE11_LITRU
3	11	31.4	5	1	B10A_CITPR
4	11	31.4	5	1	RE31_LITRU
5	11	31.4	5	1	RE32_LITRU
6	10	28.6	5	1	E104_LITRU
7	9	25.7	7	1	CCF1_ENTFA
8	9	25.7	7	1	C1A_ENTFA
9	8	22.9	5	1	ALL4_CARMA
10	8	22.9	5	1	PSK_DAVCA
11	8	22.9	5	1	RE21_LITRU
12	8	22.9	7	1	ALL2_CARMA
13	8	22.9	7	1	ALL3_CARMA
14	8	22.9	7	1	ALL4_CARMA
15	8	22.9	7	1	ALL5_CARMA
16	8	22.9	7	1	FAR1_ASCSU
17	7	20.0	4	1	RM01_YEAST
18	7	20.0	6	1	FARP_MONEX
19	7	20.0	6	1	LOK1_LOCM1
20	7	20.0	6	1	UN06_CLOPA
21	7	20.0	7	1	FAR1_PROCL
22	7	20.0	7	1	FAR2_PROCL
23	7	20.0	7	1	FAR4_PANRE
24	7	20.0	7	1	FARB_CALVO
25	7	20.0	7	1	GFRP_MOUSE
26	7	20.0	7	1	IGAO_DACDE
27	6	17.1	3	1	LUXE_VIBET
28	6	17.1	4	1	ACH1_ACHHU
29	6	17.1	4	1	FAR3_HIRME
30	6	17.1	4	1	FAR4_HIRME
31	6	17.1	4	1	FLRF_HIRME
32	6	17.1	4	1	FLRF_HIRME
33	6	17.1	4	1	FLRN_ANTEL

34	6	17.1	4	1	FMRE_MACNI	P01162 macrocallis
35	6	17.1	4	1	FRYI_ANTEL	P58706 anthopleura
36	6	17.1	4	1	OCPI_OCTMI	P58649 octopus min
37	6	17.1	4	1	OCPI_OCTMI	P58649 octopus min
38	6	17.1	5	1	E103_LITRU	P82099 litoria rub
39	6	17.1	5	1	FARP_ARTTR	P41853 artloposthi
40	6	17.1	5	1	PAP2_PARMA	P1864 pardachirus
41	6	17.1	5	1	SUGA_ACHDO	P19991 acheta dome
42	6	17.1	5	1	TPIS_CANPA	P54714 canis fam11
43	6	17.1	5	1	TRM3_ECOLI	P13973 escherichia
44	6	17.1	5	1	UC22_MAIZE	P80628 zea mays (m
45	6	17.1	5	1	UXA4_CHLTR	P38005 chlamydia t

ALIGNMENTS

RESULT 1
ALL7_CYPDPO STANDARD: PRT: 7 AA.
AC P82158;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 7.
OS Cydia pomonella (Coddling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylsia;
OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Duvé H., Johnsen A.H., Maestri J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 873 MW; 672879CAB569350 CRC64;

Query Match 40.0%; Score 14; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1e+03; 1; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 DFTL 5
DB 4 DFTL 7
RESULT 2
ID RE11_LITRU STANDARD: PRT: 5 AA.
AC P82070;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidina 1.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinbörner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
'Litoria rubella'. the skin peptide profile as a probe for the study

RT of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=598; METHOD=FAB.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA: 598 MW: 6DD9C9CAB2A00000 CRC64;

Query Match 34.3%; Score 12; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
 Db 2 DF 3

RESULT 3
 BIOA_CITFR STANDARD; PRT; 5 AA.
 AC P13071;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Adenosylmethionine-8-amino-7-oxononanoate aminotransferase
 DE (EC 2.6.1.62) (7,8-diamino-pelargonic acid aminotransferase) (DAPA
 DE aminotransferase) (Fragment).
 GN BIOA.
 OS Citrobacter freundii.
 OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae;
 OC Citrobacter.
 OX NCBI_Taxid=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89006280; PubMed=2971595;
 RA Shihuan D., Campbell A.;
 RT "transcriptional regulation and gene arrangement of Escherichia coli,
 RT Citrobacter freundii and Salmonella typhimurium biotin operons.";
 RL Gene 67:203-211(1988).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 8-amino-7-oxononanoate -> S-adenosyl-4-methylthio-2-oxobutanoate + 7,8-diaminononanoate.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: BIOTIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
 CC -----
 CC EMBL: M21922; -; NOT_ANNOTATED_CDS.
 DR InterPro: IPR000954; AminoTran_3.
 DR PROSITE: PS00600; AA_TRANSFERRER_CLASS_3; PARTIAL.
 KW Biotin biosynthesis; Transferase; Aminotransferase;
 KW pyridoxal phosphate.
 FT NON_TER 5
 SQ SEQUENCE 5 AA: 582 MW: 6AABAB1BA6F00000 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TD 2
 Db 1

Db 3 TD 4

RESULT 4
 RE33_LITRU STANDARD; PRT; 5 AA.
 AC P82072;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 3.1.
 OS Litoria rubella (Desert tree frog)
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_Taxid=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC Tissue-skin secretion;
 RA Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
 RA Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 RT 'Litoria rubella', the skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAB.
 KW Amphibian skin; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA: 656 MW: 71A9C9CB10300000 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FT 4
 Db 4 FT 5

RESULT 5
 RE32_LITRU STANDARD; PRT; 5 AA.
 AC P82073;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 3.2.
 OS Litoria rubella (Desert tree frog).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_Taxid=104895;
 RN [1]
 RP SEQUENCE.
 RC Tissue-skin secretion;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria rubella. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA: 570 MW: 71A9C9CB82A00000 CRC64;

Query Match 31.4%; Score 11; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FT 4
11
DB 4 FT 5

RESULT 6
EID4_LITRU STANDARD; PRT; 5 AA.
AC P82100;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Electric 4.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE-Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RA "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:0-0(1999).
CC -1- SUPRACELLULAR LOCATION: SECRETED.
KW Amphibian skin; Amidation.
FT MOD.RES 5
SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 28.6%; Score 10; DB 1; Length 5;
Best Local Similarity 33.3%; Pred. No. 1e+05; Indels 0; Gaps 0;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 LRT 7
11
DB 2 ITV 4

RESULT 7
CCFL_ENTFA STANDARD; PRT; 7 AA.
AC P20104;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CCF10.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=89008313; PubMed=3139658;
RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA Adachi J.C., Dunny G.M., Suzuki A.;
RA "Structure of ccf10, a peptide sex pheromone which induces
RT conjugative transfer of the Streptococcus faecalis tetracycline
RT resistance plasmid, pCF10."
RL J. Biol. Chem. 263:14574-14578(1988).
CC -1- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC HEMOLYSIN PLASMID PCF10.
KW PIR: A30812; A30812.
SQ SEQUENCE 7 AA; 790 MW; 72C9D2C731B2C740 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TTL 5
11
DB 3 TTL 4

RESULT 8
CIA_ENTFA STANDARD; PRT; 7 AA.
AC P11932;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=87005252; PubMed=3093276;
RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
RA White B.A., An F.Y., Clewell D.B., Suzuki A.;
RA "Isolation and structure of the Streptococcus faecalis sex pheromone,
RT CAM373."
RL FEBS Lett. 206:69-72(1986).
CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
CC HARBORING PAM373.
CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
DR PIR: A25269; A25269.
KW Pheromone.
SQ SEQUENCE 7 AA; 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 25.7%; Score 9; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05; Indels 1; Gaps 0;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 TTL 5
11
DB 3 TTL 5

RESULT 9
AL14_CARMA STANDARD; PRT; 5 AA.
AC P81817;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas 14.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorne A.;
RA "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas."
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD.RES 5
SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 22.9% Score 8; DB 1; Length 5;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 FTL 5
 DB 3 FGL 5

RESULT 10

PSK_DAUCA STANDARD; PRT; 5 AA.
 AC PS8261;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phytosulfoline-alpha (PSK-alpha) [contains: Phytosulfoline-beta (PSK-beta)].
 OS Daucus carota (Carrot).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
 OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
 OX NCBI_TaxID=4039;
 RN [1]
 RP SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
 RC STRAIN=cv. US-Harumakigosun;
 RX MEDLINE=20121743; PubMed=10750705;
 RA Hanael H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
 RA Kamada H., Sakagami Y.;
 RT "A secreted peptide growth factor, phytosulfoline, acting as a
 RT stimulatory factor of carrot somatic embryo formation.";
 RL Plant Cell Physiol. 41:27-32(2000).
 CC -1- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
 CC CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
 CC EMBRYOS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
 CC PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOLINE FAMILY.
 KW Growth factor; Sulfation.
 FT PEPTIDE 1 4 PHYTOSULFOLINE-BETA.
 FT MOD.RES 1 1 SULFATION.
 FT MOD.RES 3 3 SULFATION.
 SQ SEQUENCE 5 AA; 687 MW; 76C1BB504B300000 CRC64;

Query Match 22.9% Score 8; DB 1; Length 5;
 Best Local Similarity 50.0%; Pred. No. 1e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTL 4
 DB 3 YL 4

RESULT 11

RE21_LITRU STANDARD; PRT; 5 AA.
 ID RE21_LITRU
 AC P82071;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Rubellidin 2.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Skin secretion;

RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
 RA Tyler M.D., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 RT 'Litoria rubella', the skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MN=626; METHOD=FAE.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

RESULT 12

ALL2_CARMA STANDARD; PRT; 7 AA.
 ID ALL2_CARMA
 AC P81805;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 2.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunodea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Dve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 RA Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 RT allatostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD.RES 7 7 AMIDATION (POTENTIAL).
 FT MOD.RES 7 7 AMIDATION (POTENTIAL).
 SQ SEQUENCE 7 AA; 770 MW; 672879CDBC5DBD70 CRC64;

OY 2 DF 3
 DB 2 EF 3

RESULT 13

ALL3_CARMA STANDARD; PRT; 7 AA.
 ID ALL3_CARMA
 AC P81806;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 3.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunodea; Portunidae; Carcinus.

OX NCB1_TaxID=6759;
 RN [1]
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Multigene family.
 SO SEQUENCE 7 AA; 796 MW; 672879CDBC476B70 CRC64;

Query Match 22.9%; Score 8; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

OY 3 FTL 5
 | |
 DB 5 FGL 7

RESULT 14
 ALL4_CARMA
 ID ALL4_CARMA STANDARD: PRT; 7 AA.
 AC P81807; 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 4.
 OS *Carcinus maenas* (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunidae; Portunidae; Carcinus.
 OX NCB1_TaxID=6759;
 RN [1]
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Multigene family.
 SO SEQUENCE 7 AA; 782 MW; 672879CDBC476AC0 CRC64;

Query Match 22.9%; Score 8; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

OY 3 FTL 5
 | |
 DB 5 FGL 7

RESULT 15
 ALL5_CARMA
 ID ALL5_CARMA STANDARD: PRT; 7 AA.
 AC P81808; 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 5.
 OS *Carcinus maenas* (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunidae; Portunidae; Carcinus.
 OX NCB1_TaxID=6759;
 RN [1]
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RX MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
 Thorpe A.;
 RT "Isolation and identification of multiple neuropeptides of the
 allatostatin superfamily in the shore crab *Carcinus maenas*.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Multigene family.
 SO SEQUENCE 7 AA; 781 MW; 672879CDBC476420 CRC64;

Query Match 22.9%; Score 8; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

OY 3 FTL 5
 | |
 DB 5 FGL 7

Search completed: July 15, 2002, 13:38:10
 Job time: 710 sec

10

11

12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:37:39 ; Search time 86.73 Seconds
(without alignments)
13.962 Million cell updates/sec

Title: US-09-712-819A-5
Perfect score: 35
Sequence: 1 TDFLUT1 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 65

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriophage:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	34.3	5	13 P82070	P82070 litorea rub
2	11	31.4	5	13 P82072	P82072 litorea rub
3	11	31.4	5	13 P82073	P82073 litorea rub
4	11	31.4	7	2 054248	054248 streptomyces
5	10	28.6	5	13 P82100	P82100 litorea rub
6	10	28.6	7	2 034028	034028 sphingomonas
7	9	25.7	7	2 007354	007354 synechococcus
8	9	25.7	7	10 P82445	P82445 nitroclitella
9	8	22.9	5	13 P82071	P82071 litorea rub
10	8	22.9	7	2 P70804	P70804 azotobacter
11	7	20.0	5	2 P83073	P83073 bacillus ce
12	7	20.0	7	2 047029	047029 enterobacter
13	7	20.0	7	2 P72081	P72081 nocardia ja
14	7	20.0	7	8 098866	098866 spinacia ol
15	6	17.1	5	13 P82099	P82099 litorea rub
16	6	17.1	6	10 P82541	P82541 spinacia ol

17	6	17.1	6	13 P82096	P82096 litorea rub
18	6	17.1	7	2 050556	050556 actinobacill
19	6	17.1	7	4 015903	015903 homo sapien
20	6	17.1	7	6 028742	028742 oryctolagus
21	6	17.1	7	10 049223	049223 glycine max
22	6	17.1	7	10 09C5B3	09C5B3 arabidopsis
23	6	17.1	7	11 063480	063480 rattus norv
24	6	17.1	7	11 055184	055184 rattus norv
25	6	17.1	7	12 09Y010	09Y010 transmisisib
26	6	17.1	7	13 P82065	P82065 litorea rub
27	5	14.3	6	10 P82181	P82181 spinacia ol
28	5	14.3	6	10 P82182	P82182 spinacia ol
29	5	14.3	7	2 047505	047505 escherichia
30	5	14.3	7	2 095945	095945 saccharomyc
31	5	14.3	7	10 P93233	P93233 lycopersico
32	5	14.3	7	12 067113	067113 influenza a
33	5	14.3	7	12 065578	065578 bovine herp
34	5	14.3	7	13 042564	042564 fuqu rubrip
35	5	14.3	7	15 007624	007624 rous sarcom
36	4	11.4	4	11 008433	008433 rattus norv
37	4	11.4	7	4 015897	015897 homo sapien
38	4	11.4	7	8 P92421	P92421 psathyrosta
39	4	11.4	7	8 P92385	P92385 hordeum mar
40	4	11.4	7	8 P92372	P92372 haynaldia v
41	4	11.4	7	8 P92403	P92403 lophopyrum
42	4	11.4	7	8 P92425	P92425 pseudoroegn
43	4	11.4	7	8 P92387	P92387 nematoda p
44	4	11.4	7	8 P92427	P92427 peridictyon
45	4	11.4	7	8 P92390	P92390 heterantheon

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	5 AA.
P82070	P82070			
AC	P82070:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel. 13, Last annotation update)			
DE	RUBELLIDIN 1.1.			
OS	litorea rubella (Desert tree frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;			
OC	litorea.			
OX	NCBI_TaxID=104895;			
RN	[1]			
RP	SEQUENCE AND MASS SPECTROMETRY.			
RC	TISSUE-SKIN SECRETION.			
RA	Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,			
RA	Tyler M.J., Wallace J.C.,			
RT	"The structure of new peptides from the Australian red tree frog			
RT	'litorea rubella', the skin peptide profile as a probe for the study			
RT	of evolutionary trends of amphibians.";			
RL	Aust. J. Chem. 49:955-963(1996).			
CC	- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR			
CC	ANTIBIOTIC ACTIVITY.			
CC	- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.			
CC	- MASS SPECTROMETRY: MW=598; METHOD=FAH.			
KW	Amphibian skin.			
SO	SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;			

Query Match 34.3%; Score 12; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
 P82072 PRELIMINARY; PRT: 5 AA.
 AC P82072; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE RUBELLIDIN 3.1.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE AND MASS SPECTROMETRY.
 RC TISSUE=SKIN SECRETION;
 RA Steinboerner S.T., Wabnitz P.A., Maugh R.J., Bowle J.H., Gao C.,
 Tyler M.J., Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 'Litoria rubella', the skin peptide profile as a probe for the study
 of evolutionary trends of amphibians.";
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTI-BIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MW=655; METHOD=FA/AB.
 KW Amphibian skin; Amidation.
 FT MOD.RES 5
 SQ SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 31.4%; Score 11; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FT 4
 Db 4 FT 5

RESULT 3
 P82073 PRELIMINARY; PRT: 5 AA.
 AC P82073; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE RUBELLIDIN 3.2.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN SECRETION;
 RA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 'Litoria rubella', comparison with the skin peptides from Litoria
 rubella.";
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTI-BIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin.
 FT MOD.RES 5
 SQ SEQUENCE 5 AA; 570 MW; 71A9C9CB62A00000 CRC64;

Query Match 31.4%; Score 11; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FT 4

Db 4 FT 5
 RESULT 4
 054248 PRELIMINARY; PRT: 7 AA.
 ID 054248; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE RPLO PROTEIN (FRAGMENT).
 GN RPLO.
 OS Streptomyces griseus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 CC NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N2-3-11;
 RX MEDLINE=20011291; PubMed=10542330;
 RA Poehling S., Piepersberg W., Wehmeyer U.F.;
 RT "Analysis and regulation of the sec Y gene from Streptomyces griseus
 N2-3-11 and interaction of the SecY protein with the SecE protein.";
 RL Biochim. Biophys. Acta 1447:298-302(1999).
 DR EMBL; X59515; CAA65160.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 31.4%; Score 11; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TLT 6
 Db 2 TLT 4

RESULT 5
 P82100 PRELIMINARY; PRT: 5 AA.
 ID P82100; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE ELECTRIN 4.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 CC Litoria.
 RX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN SECRETION;
 RA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 'Litoria rubella', comparison with the skin peptides from Litoria
 rubella.";
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTI-BIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin; Amidation.
 FT MOD.RES 5
 SQ SEQUENCE 5 AA; 616 MW; 61F2D1A059A00000 CRC64;

Query Match 28.6%; Score 10; DB 13; Length 5;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 LTI 7
 Db 2 LTI 4

RESULT 6
 ID 034028 PRELIMINARY; PRT; 7 AA.
 AC 034028;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CATECHOL-2,3-DIOXYGENASE (FRAGMENT).
 GN PHNE.
 OS Sphingomonas chungbukensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;
 CC Sphingomonas.
 OX NCBI_TaxID=56193;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DJ77;
 RA Kim Y.-C.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88298; AAB6311.1; -
 KW Dioxigenase.
 FT NON_TER
 SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;

Query Match 28.6%; Score 10; DB 2; Length 7;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 LTI 7
 : :
 DB 3 MTV 5

RESULT 7
 ID 007354 PRELIMINARY; PRT; 7 AA.
 AC 007354;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE NIFK (FRAGMENT).
 GN NIFK.
 OS Synechococcus sp. (Strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
 OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
 OX NCBI_TaxID=41431;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RF-1;
 RA MEDLINE=9231861; PubMed=10217509;
 RA Huang T.C.; Lin R.F.; Chu M.K.; Chen H.M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RT RF-1."
 RL Microbiology 145:743-753(1999).
 DR EMBL; AF003700; AAC35193.1; -
 FT NON_TER
 SQ SEQUENCE 7 AA; 849 MW; 7412C72A9D5B030 CRC64;

Query Match 25.7%; Score 9; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 FTL 5
 : :
 DB 3 FDL 5

RESULT 8
 ID P82445 PRELIMINARY; PRT; 7 AA.
 AC P82445;

DT 01-JUN-2000 (TREMBlrel. 14, Created)
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE 10 KDA CELL WALL PROTEIN (FRAGMENT).
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE.
 RC STRAIN-CV. PETIT HAVANA;
 RA Blee K.A.; Bonham V.A.; Mitchell G.P.; Robertson D.; Stabas A.R.;
 RA Wojtaszek P.; Bolwell G.P.;
 RT "Proteomic study of secondary cell wall proteins from transformed
 RT tobacco culture."
 RL Planta 0:0-0(2000).
 CC -1- SURCELLULAR LOCATION: CELL WALL.
 CC -1- TISSUE SPECIFICITY: XYLEM.
 KW Cell wall.
 FT NON_TER
 SQ SEQUENCE 7 AA; 758 MW; 69D2C1E862D1B2A0 CRC64;

Query Match 25.7%; Score 9; DB 10; Length 7;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 LTI 7
 : :
 DB 1 VTV 3

RESULT 9
 ID P82071 PRELIMINARY; PRT; 5 AA.
 AC P82071;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE RUBELLIDIN 2.1.
 OS Litorea rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
 OC Litorea.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE-SKIN SECRETION;
 RA Steinboerner S.T.; Wabnitz P.A.; Waugh R.J.; Bowie J.H.; Gao C.;
 RA Tyler M.J.; Wallace J.C.;
 RT "The structure of new peptides from the Australian red tree frog
 RT 'Litorea rubella', the skin peptide profile as a probe for the study
 RT of evolutionary trends of amphibians."
 RL Aust. J. Chem. 49:955-963(1996).
 CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
 CC ANTIBIOTIC ACTIVITY.
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -1- MASS SPECTROMETRY: MM=626; METHOD=FAB.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA; 626 MW; 6DD9C9CB10300000 CRC64;

Query Match 22.9%; Score 8; DB 13; Length 5;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
 : :
 DB 2 EF 3

RESULT 10

P70804
 ID P70804 PRELIMINARY; PRT; 7 AA.
 AC P70804;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALG1 PROTEIN (FRAGMENT).
 GN ALG1.
 OS Acetobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Acetobacter.
 RX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E;
 RX MEDLINE=96427318; PubMed=8830682;
 RA Rehm B.H.A., Ertesvag H., Valla S.;
 RT "A new Acetobacter vinelandii mannuronan C-5-epimerase gene (algC) is
 RT part of an alg gene cluster physically organized in a manner similar
 RT to that in Pseudomonas aeruginosa.";
 RL J. Bacteriol. 178:5884-5889(1996).
 DR EMBL: X87973; CAA61230.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A2D1AED0 CRC64;

Query Match 22.9%; Score 8; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0;

QY 6 TI 7
 1;
 Db 2 TV 3

RESULT 11
 P83073
 ID P83073 PRELIMINARY; PRT; 5 AA.
 AC P83073;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE 88 KDA PROTEIN (FRAGMENT).
 OS Bacillus cereus.
 OC Bacillus; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1396;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=NCIMB 11796;
 RA Browne N., Dowds B.C.A.;
 RL Submitted (Jul-2001) to the SWISS-PROT data bank.
 FT NON_TER 5
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F00000 CRC64;

Query Match 20.0%; Score 7; DB 2; Length 5;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0;

QY 1 TD 2
 1;
 Db 4 TE 5

RESULT 12
 Q47029
 ID Q47029 PRELIMINARY; PRT; 7 AA.
 AC Q47029;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE AAD A1 PROTEIN (FRAGMENT).

GN AAD A1.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Enterobacter.
 OX NCBI_TaxID=550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94079349; PubMed=8257126;
 RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
 RT "Analysis of the dac(3)-Via gene encoding a novel 3-N-
 RT acetyltransferase.";
 RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
 DR EMBL: M88012; AAA16193.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 20.0%; Score 7; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0;

QY 5 LT 6
 1;
 Db 1 IT 2

RESULT 13
 P72081
 ID P72081 PRELIMINARY; PRT; 7 AA.
 AC P72081;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 3'-METHYLCEPHEM HYDROXYLASE (FRAGMENT).
 GN CEF3.
 OS Nocardia lactamdurans.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Pseudonocardineae; Pseudonocardiaceae; Amycolatopsia.
 OX NCBI_TaxID=1913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96009872; PubMed=7557411;
 RA Coque J., Perez-Llarena F.J., Enguita F.J., Fuente J.L., Martin J.F.,
 RA Liras P.;
 RT "Characterization of the cmcH genes of Nocardia lactamdurans and
 RT Streptomyces clavuligerus encoding a functional 3'-hydroxymethylcephem
 RT O-carbamoyltransferase for cephamycin biosynthesis.";
 RL Gene 162:21-27(1995).
 DR EMBL: Z21682; CAA79797.1; -;
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 746 MW; 75A1B2CDD1E681C0 CRC64;

Query Match 20.0%; Score 7; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0;

QY 4 TL 5
 1;
 Db 1 TM 2

RESULT 14
 O98866
 ID O98866 PRELIMINARY; PRT; 7 AA.
 AC O98866;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE CYTOCHROME B/F SUBUNIT IV (FRAGMENT).
 OS Spinacia oleracea (Spinach).
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86120353; Pubmed=3003688;
 RA Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
 RT "Spinach plastid genes coding for initiation factor IF-1, ribosomal
 RT protein S11 and RNA polymerase alpha-subunit.";
 RU Nucleic Acids Res. 14:1029-1044(1986).
 DR EMBL: X03496; CAA27215.1; -.
 KW Chloroplast.
 FT NON_TER
 SQ SEQUENCE 7 AA; 907 MW; 644729D77409C420 CRC64;

Query Match 20.0%; Score 7; DB 8; length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0;

OY 2 DF 3
 : 1
 Db 1 NF 2

RESULT 15
 P82099
 ID P82099 PRELIMINARY; PRT; 5 AA.
 AC P82099;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 DE ELECTRIN 3.
 OS Litoria rubella (Desert tree frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidea; Hylidae;
 OC Litoria.
 OX NCBI_TaxID=104895;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=SKIN SECRETION;
 RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litori electrica. Comparison with the skin peptides from Litoria
 RT rubella.";
 RT Aust. J. Chem. 52:0-0(1999).
 KW Amphibian skin; Amidation.
 FT MOD_RES 5
 SQ SEQUENCE 5 AA; 630 MW; 668761F2C9A00000 CRC64;

Query Match 17.1%; Score 6; DB 13; length 5;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 0; Mismatches 0; Indels 0;

OY 3 F 3
 : 1
 Db 1 F 1

Search completed: July 15, 2002, 13:37:39
 Job time: 729 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:25:26 ; Search time 107.71 Seconds
(without alignments)
7.219 Million cell updates/sec

Title: US-09-712-819A-5

Perfect score: 35

Sequence: 1 TDFRTLI 7

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues 52936

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing:

Minimum Match 0%

Maximum Match 100%

Database :

A_Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	30	85.7	7 20	AAV40738
2	30	85.7	7 21	AAAB30076
3	27	77.1	7 20	AAV40737
4	27	77.1	7 21	AAAB30075
5	26	74.3	7 20	AAV40736
6	26	74.3	7 21	AAAB30074
7	21	60.0	7 20	AAV14399
8	20	57.1	6 19	AAW31467
9	19	54.3	5 9	AAAB2200
10	19	54.3	6 20	AAV06466
11	19	54.3	6 21	AAV77708

12	19	54.3	6 22	AAE13076	Epitope tag #4 use
13	19	54.3	6 22	AAAS1422	Integrin activatin
14	19	54.3	6 22	AAAB97355	AU5 epitope used 1
15	19	54.3	6 22	AAAB59859	AU5 peptide epitop
16	19	54.3	7 21	AAV52584	Amaranthus viridis
17	18	51.4	6 19	AAW75358	Hexapeptide #13 bl
18	18	51.4	6 19	AAW75290	Hexapeptide #13 bl
19	18	51.4	7 11	AAK09440	LRA-1 alpha subunl
20	18	51.4	7 19	AAAS8711	Tryptic 40 kD subu
21	18	51.4	7 20	AAV14403	Peptide CDR-H1-7 d
22	18	51.4	7 22	AAAG44484	Antihpatitis C pe
23	17	48.6	5 21	AAAS2195	Human anti-HBS ant
24	17	48.6	5 21	AAAB17215	IL-1 antagonist pe
25	17	48.6	5 21	AAV51466	AAV VP3 derived pe
26	17	48.6	6 20	AAV33711	Hepatoma diagnosi
27	17	48.6	6 20	AAK06532	Epidermal growth I
28	17	48.6	6 21	AAV95333	Human pancreatic p
29	17	48.6	6 22	AAAG8217	Human SNP associat
30	17	48.6	6 22	AAAB87699	Hepatoma diagnosi
31	17	48.6	6 22	AAAS5567	T cell surface rec
32	17	48.6	6 22	AAAS7414	CD90 C-terminal co
33	17	48.6	6 22	AAAS7824	CD90 C-terminal co
34	17	48.6	7 11	AAK07656	Ribonuclease reduc
35	17	48.6	7 13	AAK48974	Human betat.6-N-ac
36	17	48.6	7 16	AAW11972	T-cell epitope #3
37	17	48.6	7 16	AAAR2752	Antimalarial pepti
38	17	48.6	7 16	AAAR2753	Antimalarial pepti
39	17	48.6	7 16	AAAR2754	Antimalarial pepti
40	17	48.6	7 16	AAAR2755	Antimalarial pepti
41	17	48.6	7 20	AAV40723	S4 derivative #20
42	17	48.6	7 20	AAV40735	S4 derivative #9,
43	17	48.6	7 20	AAV07710	Liquid Interferon
44	17	48.6	7 21	AAAB30062	Scaffold protein S
45	17	48.6	7 21	AAAB30073	Scaffold protein S

ALIGNMENTS

RESULT 1	
AAV40738	standard; peptide: 7 AA.
ID	AAV40738 standard; peptide: 7 AA.
AC	AAV40738;
DT	01-DEC-1999 (first entry)
DE	S4 derivative #12, beta strand of scaffold protein structure.
KW	Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW	tumour; chemotherapeutic agent.
OS	Synthetic.
XX	
PN	EP947582-A1.
XX	
PD	06-OCT-1999.
XX	
PF	31-MAR-1998; 98EP-0870065.
XX	
PR	31-MAR-1998; 98EP-0870065.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX	
DR	WPI; 1999-542958/46.
XX	
PT	New scaffold protein, useful for stabilizing antigens used as vaccines
XX	
PS	Disclosure: Page 6; 105pp; English.
XX	

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a beta fold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separable molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

Query Match	85.7%;	Score 30;	DB 20;	Length 7;
Best Local Similarity	100.0%;	Pred. No. 6.4e+05;		
Matches	6;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps	0

QY	2	DEFTLTI	7
Db	1	dfctltl	6

RESULT	2
AAB30076	
ID	AAB30076 standard; Peptide; 7 AA

AA	AAB30076;
AC	
XX	
DT	09-FEB-2001 (first entry)

DE Scaffold protein SCA 54 peptide SEQ ID NO: 137.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis,
KW diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN W0200060070-A1.

PD 12-OCT-2000

PF 01-APR-1999; 99WO-EP02283

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI; 2000-665002/64

Scaffold composed of single-chain polypeptide having beta sandwich
architecture carrying new and randomized peptide sequences useful as
supporting framework and carrying antigen- or receptor binding
fragments -

Dislosure; Page 15; 68pp; English.

PS Disclosure; page 15; 68pp; English.

xx The present invention is concerned with producing scaffold proteins based upon the human CTRP4-SCA domain. These scaffold proteins can be used as a scaffold to bind antigen- or receptor-binding fragments. These can be used in the treatment of diseases such as cancer, atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and diabetic retinopathy. Sequences AA32930-329939 were used in the production of the proteins of the invention.

Sequence 7 AA;

Query Match	85.7%;	Score 30;	DB 21;	Length 7;
Best Local Similarity	100.0%;	Pred. No. 6,4e+05;		
Matches	6;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

Qy	2	DEFTLI	7
Db	1	dftltl	6

RESULT	3
AAV40737	
ID	AAV40737 standard; peptide; 7 AA.

AA	AA
AC	AA
XX	AA
DT	AA

DE S4 derivative #11, beta strand of scaffold protein structure.

AA Scaffold protein; beta strand; stabilize antigen; vaccine,
 KW tumour; chemotherapeutic agent.

Synthetic.

PN EP947582-A1.

PD 06-OCT-1999

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI; 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure; Page 6; 105pp; English.

CC sequenced AAAY0727-V407748 are functionally equivalent derivatives of the
CC S4 peptide (AAAY0607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAAY0601-V40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure
CC the scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will

target the complex to tumour cells. Another surface may be bound to a cytotoxic molecule or an autoimmune antibody which may then kill the tumour cells. Therefore the scaffold protein may be used to target chemotherapeutic agents to specific cells. It may also be used to stabilize individual peptides in a peptide library and may be used in diagnostic techniques, and to stabilize antigens used as vaccines.

Sequence 7 AA:

Query Match 77.1%; Score 27; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLTI 7
1:|||||
Db 1 dftlti 6

RESULT 4

AAB30075
ID AAB30075 standard; Peptide; 7 AA.

AC AAB30075;

DT 09-FEB-2001 (first entry)

XX Scaffold protein SCA S4 peptide SRQ ID NO: 136.

XX Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

XX Synthetic.

XX WO200060070-A1.

XX 12-OCT-2000.

XX 01-APR-1999; 99WO-EP02283.

XX 01-APR-1999; 99WO-EP02283.

XX (INNO-) INNOGENETICS NV.

XX Desmet J, Hufton S, Hoogenboom H, Sablon E;

XX WPI; 2000-665002/64.

XX Scaffold composed of single-chain polypeptide having beta sandwich

XX architecture carrying new and randomized peptide sequences useful as

XX supporting framework and carrying antigen- or receptor binding

XX fragments -

XX Disclosure; Page 15; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
XX based upon the human CTLA-4 SCA domain. These scaffold proteins can be
XX used as a scaffold to bind antigen- or receptor-binding fragments. These
XX can be used in the treatment of diseases such as cancer,
XX atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
XX CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
XX production of the proteins of the invention.

XX Sequence 7 AA:

Query Match 77.1%; Score 27; DB 21; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLTI 7
1:|||||

Db 1 dftlti 6

RESULT 5

AAV40736
ID AAV40736 standard; peptide; 7 AA.

XX AAV40736;

XX 01-DEC-1999 (first entry)

XX S4 derivative #10, beta strand of scaffold protein structure.

XX Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.

XX Synthetic.

XX EP947582-A1.

XX 06-OCT-1999.

XX 31-MAR-1998; 98EP-0870065.

XX 31-MAR-1998; 98EP-0870065.

XX (INNO-) INNOGENETICS NV.

XX Desmet J, Hufton S, Hoogenboom H, Sablon E;

XX WPI; 1999-542958/46.

XX New scaffold protein, useful for stabilizing antigens used as vaccines

XX -

XX Disclosure; Page 6; 105pp; English.

XX Sequences AAV40727-Y40748 are functionally equivalent derivatives of the
XX S4 peptide (AAV40607) which forms part of a scaffold protein. S4 is a
XX beta strand peptide which forms part of a beta sheet. Peptides
XX (AAV40601-Y40609) together form a single-chain scaffold protein which
XX contains at least 1 disulfide bond, contains less than 10% alpha helix
XX and contains at least 6 beta-strands. The scaffold protein is constructed
XX of beta strands S1-S6, and may also include these sequences. The beta strands
XX functionally equivalent derivative of these sequences. The beta strands
XX form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
XX the next by hydrogen bonds, which generate a beta sandwich architecture.
XX If the additional beta strands A1-A3 are included in the structure the
XX scaffold is constructed of two beta sheets, with the structures
XX A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
XX other via amino acid loops, where at least one of the loops binds to a
XX receptor or antigen. The scaffold protein is used to stabilize antigens
XX or whole proteins such as receptors, or their fragments. It may be used
XX to bind two separate molecules. For example, one surface of the scaffold
XX may be bound to a protein which binds to a tumour antigen. This will
XX target the complex to tumour cells. Another surface may be bound to a
XX cytotoxic molecule or an autoimmune antibody which may then kill the
XX tumour cells. Therefore the scaffold protein may be used to target
XX chemotherapeutic agents to specific cells. It may also be used to
XX stabilize individual peptides in a peptide library and may be used in
XX diagnostic techniques, and to stabilize antigens used as vaccines.

XX Sequence 7 AA:

Query Match 74.3%; Score 26; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLTI 7
1:|||||
Db 1 dftlti 6

RESULT 6
AAB30074
ID AAB30074 standard; Peptide: 7 AA.
XX
AC AAB30074;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA s4 peptide SEQ ID NO: 135.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PE 01-APR-1999; 99WO-EP02283.
XX
PR 01-APR-1999; 99WO-EP02283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufon S, Hoogenboom H, Sablon E;
XX
DR WPI: 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments -
XX
PS Disclosure; Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.
XX
SQ Sequence 7 AA;

Query Match 74.3%; Score 26; DB 21; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFFLT 7
 |||||
 1 dffl 6
Db 1 dffl 6

RESULT 7
AAV14399
ID AAV14399 standard; peptide: 7 AA.
XX
AC AAV14399;
XX
DT 17-AUG-1999 (first entry)
XX
DE Peptide CDR-H1-7(Y3F) derived from anti-HCV protease MAb 8D4.
XX
KW Complementarity determining region; CDR: monoclonal antibody; MAb;
KW hepatitis C virus; HCV; protease; binding site.
XX
OS Synthetic.
XX

PN JP11127861-A.
XX
PD 18-MAY-1999.
XX
PE 29-OCT-1997; 97JP-0297451.
XX
PR 29-OCT-1997; 97JP-0297451.
XX
PA (NIHA) JAPAN ENERGY CORP.
XX
DR WPI: 1999-350322/30.
XX
PT Neutralized antibody partial peptide derived from hepatitis C virus
PT - useful for inhibiting Hepatitis C virus (HCV) serine protease
PT activity
XX
PS Example 1; Page 24; 32pp; Japanese.
XX
CC This sequence corresponds to a peptide (CDR-H1-7; AAV14403) derived from
CC the sequence of the heavy chain variable region complementarity
CC determining region (CDR)-1 of the anti-hepatitis C virus (HCV) Ser/Thr
CC protease monoclonal antibody (MAb) 8D4 protein. The peptide has a Tyr
CC to Phe amino acid substitution at position 3 compared to the CDR-H1-7
CC peptide. The invention relates to the use of partial peptides
CC (AAV14348-Y14353) from the MAb 8D4 for inhibiting HCV serine protease
CC activity.
XX
SQ Sequence 7 AA;

Query Match 60.0%; Score 21; DB 20; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFLT 5
 |||||
 1 tdfvl 5
Db 1 tdfvl 5

RESULT 8
AAW31467
ID AAW31467 standard; Protein; 6 AA.
XX
AC AAW31467;
XX
DT 04-AUG-1998 (first entry)
XX
DE Transcriptional activator peptide fragment LS130.
XX
KW Activating sequence; Gal4; transcriptional activator; RNA polymerase;
KW Protein-protein interaction; gene therapy; therapeutic; holoenzyme;
KW Gal11; DNA binding domain.
XX
OS Synthetic.
XX
PN WO9744447-A2.
XX
PD 27-NOV-1997.
XX
PE 02-MAY-1997; 97WO-US07338.
XX
PR 01-MAY-1997; 97US-0017016.
XX
PR 03-MAY-1996; 96US-0017016.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Lu X, Plashne M, Wu Y;
XX
DR WPI: 1998-018502/02.
XX
DR N-PSDB: AAV02565.
XX
PT New transcriptional activator containing DNA binding domain bound to
PT peptide - useful for controlling gene expression, especially in gene

therapy, and in protein-protein interaction assays, does not inhibit other transcription activators

Example 1; Page 26; 55pp; English.

AAW31412-W31459, AAW31461-W31481, AAW31483-W31493 and AAW39053-W39076 are fragments used in an assay to determine novel transcriptional activators. The method involves the production of transcriptional activators comprising of a DNA-binding group and a 6-25 amino acid peptide that is covalently bonded to the DNA binding group and does not represent a fragment of a natural transcription activator. CC Protein-protein interactions are identified in the assay by fusing a DNA-binding domain to a library of DNA fragments and introducing this and a fusion of target protein and a polypeptide containing a region of Gal4 which interacts with Gal1p into a cell containing Gal1p and identifying members of the library that interact with the target from activation of CC transcription. Such constructs are used to activate transcription in a cell, e.g. for controlling gene activity, particularly in gene therapy CC (e.g. recognizing a site close to a selected therapeutic gene). CC Transcription can be activated without blocking other transcriptional activators. They probably act by interacting with a component of the RNA CC polymerase II holoenzyme, Gal11, the strongest known yeast activator, CC which provides a more sensitive assay allowing detection of even weak CC protein-protein interactions. Such activators do not create toxicity CC problems even when overexpressed.

Sequence 6 AA:

Query Match 57.1%; Score 20; DB 19; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFTL 5
|||
Db 1 tdfll 5

RESULT 9
AAP82200
ID AAP82200 standard; protein; 5 AA.

AAP82200;

25-OCT-1990 (first entry)

DE Example of peptide 2 for treatment of schizophrenia or psoriasis.

KM schizophrenia; psoriasis; vasoactive intestinal polypeptide.

OS synthetic.

PN SE8700125-A.

PD 16-JUL-1988.

PF 15-JAN-1987; 87SE-0000125.

PR 15-JAN-1987; 87SE-0000125.

PA (WETT/) WETTERBERG.

PI Wetterberg L;

DR WPI; 1988-328337/46.

PT Short peptides for treatment of psoriasis and schizophrenia -

PT comprise vasoactive intestinal polypeptide, peptide T or short

PS peptide with five aminoacids

CC Claim 1; Page 5; 9pp; Swedish.
Specific example of pentapeptide of the general formula of AAP82197.

CC These peptides can be administered intravenously,
CC topically or perorally to relieve the symptoms of psoriasis or
CC schizophrenia. Amino acids at posns 2 and 3 can be any residue but
CC Asp is preferred at posn 3. See also AAP82196-9.

Sequence 5 AA:

Query Match 54.3%; Score 19; DB 9; Length 5;
Best Local Similarity 75.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDFTL 4
|||
Db 2 tdyt 5

RESULT 10
AAV06466
ID AAV06466 standard; Peptide; 6 AA.

AAV06466;

27-SEP-1999 (first entry)

DE Epitope tag.

KM Epitope tag; antibody engineering; yeast; surface display;

KW protein library; peptide library.

OS Synthetic.

PN WO9336569-A1.

PD 22-JUL-1999.

PF 20-JAN-1999; 99WO-US01188.

PR 26-AUG-1998; 98US-0140084.

PR 20-JAN-1998; 98US-0005388.

PA (UNII) UNIV ILLINOIS FOUND.

PI Boder ET, Kieke MC, Kranz DM, Shusta E, Wiltrop KD;

DR WPI; 1999-430619/36.

PT Selecting proteins with enhanced phenotypic properties than

PT wild-type proteins, is useful for highly specific cancer diagnosis

PT and therapy

PS Disclosure; Page 7; 116pp; English.

CC This peptide comprises an epitope tag that can be used in
CC methods of the invention. The invention discloses a powerful new
CC system for engineering antibody affinity and specificity, by
CC constructing a microbial analogue of the mammalian system's B cell
CC repertoire. Antibodies are displayed on the surface of yeast cells
CC by genetic fusion with yeast cell wall proteins, especially
CC agglutinin proteins. After mutation, variants are selected on the
CC basis of improved binding characteristics with fluorescently
CC labeled targets. The selection method also identifies proteins
CC with enhanced phenotypic characteristics, proteins that are
CC displayed at higher levels, proteins that are secreted at higher
CC efficiency and proteins of improved stability.

Sequence 6 AA:

Query Match 54.3%; Score 19; DB 20; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTL 5
DB 1 tdfyl 5

RESULT 11

ID AAY77708 standard; peptide; 6 AA.

AC AAY77708;

DT 12-MAY-2000 (first entry)

DE A05 peptide epitope.

KW Cell surface receptor; luminescence; protein internalization;
KM drug discovery; screening assay; epitope; A05.

OS Synthetic.

PN W0200003246-A2.

PD 20-JAN-2000.

PF 13-JUL-1999; 99WO-US15870.

PR 13-JUL-1998; 98US-0092671.

PA (CELL-) CELLOMICS INC.

PI Rubln RA, Giuliano KA, Gough A, Dunlay T;

DR WPI; 2000-171170/15.

PT Automated screening method for identifying compounds which induce cell
surface receptor internalization, useful for drug discovery -

PS Example 6; Page 67; 148pp; English.

CC The invention relates to a method for identifying compounds which
inhibit internalization of cell surface receptors. Provided are an array
of locations, each containing cells with a cell surface receptor
protein, that are treated with a test compound. The protein is
luminescently labeled or contacted with a luminescently labeled cell
before or after test compound treatment. Any luminescence produced is
converted into digital data and automatically analysed to determine if
the test compound induced the protein internalization. The novel method
is used to screen for compounds which modulate cell surface receptor
protein internalization, this can be used in drug discovery, to test
compound efficacy in living biological systems. The assay method is
automated and compact. It has high throughput and uses smaller volumes of
reagents and test compounds. Sequences AAY7704-718 represent examples of
peptide epitope tags used in the course of the invention.

CC Sequence 6 AA;

Query Match 54.3%; Score 19; DB 21; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTL 5
DB 1 tdfyl 5

RESULT 12

ID AAE13076 standard; peptide; 6 AA.

AC AAE13076;

DT 28-JAN-2002 (first entry)

XX Epitope tag #4 used in yeast cell surface display of proteins.
DE Phenotypic property; yeast; cell wall protein; epitope tag.
KM Unidentified.
XX US6300065-B1.

PN 09-OCT-2001.

PF 26-AUG-1998; 98US-0140084.

PR 31-MAY-1996; 96US-018741P.

PR 30-MAY-1997; 97US-0866398.

PR 20-JAN-1998; 98US-0009388.

PA (UNIT) UNIV ILLINOIS FOUND.

PI Kleke MC, Wittup KD, Boder ET, Kranz DM, Shusta E;

DR WPI; 2001-656236/75.

PT Selecting proteins, e.g. antibodies, with enhanced phenotypic
properties relative to those of a wild-type comprises transforming
yeast cells with a vector expressing a protein to be tested fused to a
yeast cell wall protein -

PS disclosure; Column 5; 64pp; English.

CC The present invention relates to a method for selecting proteins with
enhanced phenotypic properties relative to those of a wild-type,
CC comprises transforming yeast cells with a vector expressing a protein
to be tested fused to a yeast cell wall protein. The method is
particularly useful for selecting antibodies for improved affinity and
specificity. The present sequence is an epitope tag which is used in
yeast cell surface display of proteins.

CC Sequence 6 AA;

Query Match 54.3%; Score 19; DB 22; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTL 5
DB 1 tdfyl 5

RESULT 13

ID AAM51422 standard; peptide; 6 AA.

AC AAM51422;

DT 08-JAN-2002 (first entry)

DE Integrin activating peptide SEQ ID NO 1.

KW Integrin stimulant; vulnerary; injury healing;
KM postsurgical tissue recovery.

OS Unidentified.

PN JP2001213898-A.

PD 07-AUG-2001.

PF 31-JAN-2000; 2000JP-0022469.

PR 31-JAN-2000; 2000JP-0022469.

PA (HISM) HISAMITSU PHARM CO LTD.
XX
DR WPI: 2001-629610/73.
XX
PT An injury healing and postsurgical tissue recovering integrin
XX activating peptide -
XX
PS Claim 1: Page 3; 11pp; Japanese.
XX
CC The invention relates to novel peptides with vulnerary activity, useful
CC for injury healing and postsurgical tissue recovery by acting as an
CC integrin stimulant.
XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 22; Length 6;
Best Local Similarity 60.0%; Pred. No. 6.4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 FTLTI 7
DB 1 ycttl 5

RESULT 14
AAB97355
ID AAB97355 standard; peptide; 6 AA.
AC AAB97355;
XX
DT 15-AUG-2001 (first entry)
DE
DE AUs epitope used in dual labelled receptor construction.
XX
XX Automated measurement; cell viability; epitope tag; luminescence;
KW G-protein coupled receptor; high content screen.
XX
OS Synthetic.
XX
PN WO200135072-A2.
XX
PD 17-MAY-2001.
XX
PF 09-NOV-2000; 2000WO-US30896.
XX
PR 09-NOV-1999; 99US-0164353.
PR 18-JAN-2000; 2000US-0176504.
XX
PA (CELL-) CELLOMICS INC.
XX
PI Ghosh RN, Debiasio R, Chen Y, Bellutta P, Giuliano K, Pasley JW;
XX
DR WPI: 2001-329169/34.
XX
PT Automated measurement of cell viability, involves contacting cells with
PT luminescent reporter molecule, imaging cells to get signals, converting
PT signals into digital data and using data to measure viable cell -
XX
PS Example 6; Page 52; 155pp; English.
XX
CC This invention relates to a method for the automated measurement of cell
CC viability. The method involves contacting cells with luminescent
CC reporter molecules, imaging cells to get signals, and converting the
CC signals into digital data which can be used as a measurement of cell
CC viability. Included in the invention is a computer readable storage
CC medium comprising a programme which causes the method of the invention
CC to be activated. The method is useful for cell state identification in
CC cells. The method is also useful for drug discovery. An example of the
CC invention relates to the use of inserted sequences and their ligands for
CC high content screens incorporating dual labelled receptors. The present
CC sequence represents an epitope tag used to label one end of a G-protein
CC coupled receptor (GPCR). The intracellular and extracellular domains of

CC the GPCR are distinctly labelled so that using the method of the
CC invention the extent of internalisation of the receptor can be measured.
XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 22; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFTL 5
DB 1 tdfyl 5

RESULT 15
AAB59859
ID AAB59859 standard; peptide; 6 AA.
XX
AC AAB59859;
XX
DT 30-MAR-2001 (first entry)
DE
DE AUs peptide epitope.
XX
XX AUs peptide epitope; macromolecule trafficking; endosomal system;
KW membrane receptor internalisation.
XX
OS Unidentified.
XX
PN WO200079241-A2.
XX
PD 28-DEC-2000.
XX
XX 21-JUN-2000; 2000WO-US40260.
XX
PF 21-JUN-1999; 99US-0140143.
PR 12-JUL-1999; 99US-0352171.
PR 11-AUG-1999; 99US-0148360.
PR 13-DEC-1999; 99US-0170313.
XX
PA (CELL-) CELLOMICS INC.
XX
XX Rudin RA, Gough AH, Ghosh RN, Giuliano KA, Dunlay RT;
XX
PI WPI: 2001-091619/10.
XX
DR
XX
XX Identifying compounds modulating macromolecule trafficking through
PT endosomes, using digital data obtained by converting a luminescent
PT signal from cells contacted with the compound -
XX
PS Example 6; Page 53; 113pp; English.
XX
XX The present invention relates to an automated method for identifying
CC compounds that induce or inhibit macromolecule trafficking through an
CC endosomal system. The method comprises treating cells which possess a
CC luminescently-tagged macromolecule, with a test compound, and obtaining
CC luminescent signals from the cells. The signal is converted into digital
CC data that is used to determine if the test compound has induced or
CC inhibited the trafficking. The method can also be used to identify the
CC extent of internalisation of membrane receptors, by fusing a labelled
CC peptide epitope to the different domains of the receptor e.g. the
CC extracellular domain and intracellular domain. The present invention is
CC one such peptide epitope used in the method of the present invention.
XX
SQ Sequence 6 AA;

Query Match 54.3%; Score 19; DB 22; Length 6;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFTL 5

Db 1 111 1
tdfy 5

Search completed: July 15, 2002, 13:25:27
Job time: 1458 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:26:14 ; Search time 40.89 Seconds
(without alignments)
4.181 Million cell updates/sec

Title: US-09-712-819A-5

Perfect score: 35

Sequence: 1 TDFTLRTI 7

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 39160

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/plodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/plodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/plodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/plodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/plodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/plodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	54.3	6	4	US-09-140-084-4
2	19	54.3	7	3	US-08-916-443A-8
3	19	54.3	7	4	US-08-640-737-38
4	17	48.6	5	1	US-08-405-230-10
5	17	48.6	5	2	US-08-910-990-10
6	17	48.6	7	1	US-08-136-743B-29
7	17	48.6	7	1	US-08-136-743B-30
8	17	48.6	7	1	US-08-136-743B-31
9	17	48.6	7	1	US-08-136-743B-32
10	17	48.6	7	1	US-08-405-230-5
11	17	48.6	7	2	US-08-910-990-6
12	17	48.6	7	2	US-08-739-401A-5
13	17	48.6	7	5	PCT-US93-11703-72
14	16	45.7	5	1	US-08-180-209B-14
15	16	45.7	5	1	US-08-385-745-14
16	16	45.7	5	4	US-08-591-632-17
17	16	45.7	5	4	US-08-591-632-23
18	16	45.7	5	4	US-08-591-632-26
19	16	45.7	5	4	US-08-485-388-14
20	16	45.7	5	5	US-08-474-853-14
21	16	45.7	5	5	PCT-US94-02629-14
22	16	45.7	6	1	US-08-252-995D-7
23	16	45.7	6	2	US-08-482-228-180
24	16	45.7	6	2	US-08-834-108-7
25	16	45.7	6	3	US-08-482-528-180
26	16	45.7	7	1	US-08-136-743B-55
27	16	45.7	7	1	US-08-096-946-5

28	16	45.7	7	2	US-08-177-109A-7	Sequence 7, Appl1
29	16	45.7	7	2	US-08-687-106-7	Sequence 7, Appl1
30	16	45.7	7	3	US-09-040-216-28	Sequence 28, Appl1
31	16	45.7	7	4	US-09-173-941-52	Sequence 52, Appl1
32	16	45.7	7	5	PCT-US94-07329-5	Sequence 5, Appl1
33	15	42.9	5	1	US-08-136-743B-63	Sequence 63, Appl1
34	15	42.9	5	1	US-07-789-184-126	Sequence 126, App
35	15	42.9	5	1	US-08-475-263-126	Sequence 126, App
36	15	42.9	5	1	US-08-485-886-126	Sequence 126, App
37	15	42.9	5	2	US-08-667-001-22	Sequence 22, Appl
38	15	42.9	5	2	US-08-477-134-126	Sequence 126, App
39	15	42.9	5	2	US-08-473-489A-126	Sequence 126, App
40	15	42.9	5	3	US-08-040-216-55	Sequence 55, Appl
41	15	42.9	5	3	US-08-485-595-126	Sequence 126, App
42	15	42.9	5	3	US-08-018-760-126	Sequence 126, App
43	15	42.9	6	1	US-08-136-743B-62	Sequence 62, Appl
44	15	42.9	6	1	US-08-290-448A-41	Sequence 41, Appl
45	15	42.9	6	1	US-08-290-448A-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-09-140-084-4
Sequence 4, Application US/09140084A

Patent No. 630065

GENERAL INFORMATION:

APPLICANT: Kieke, et al.

TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof

FILE REFERENCE: D6061CIP2

CURRENT FILING DATE: 1998-08-26

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 4

LENGTH: 6

TYPE: PRT

ORGANISM: Unknown

FEATURE: OTHER INFORMATION: Description of Unknown Organism: Epitope Tag

US-09-140-084-4

Query Match 54.3%; Score 19; DB 4; Length 6;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFTL 5
DB 1 TDFTL 5

RESULT 2
US-08-916-443A-8
Sequence 8, Application US/08916443A

Patent No. 6001986

GENERAL INFORMATION:

APPLICANT: Yong Sig KIM

APPLICANT: Sun Chung PARK

APPLICANT: Soo Kyung OH

APPLICANT: Hosu Lee

APPLICANT: Jeong Woo CHO

TITLE OF INVENTION: Antiviral Proteins, Amaranth 1 and 2, from

TITLE OF INVENTION: Amaranthus Viridis, DNAs Encoding Therefrom

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kenyon & Kenyon

STREET: 1025 Connecticut Avenue, N.W., Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3+ Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: Wordperfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916,443A
FILING DATE: 22 AUG 1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 1942/18
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-916-443a-8

Query Match 54.3% Score 19: DB 3: Length 7;
Best Local Similarity 50.0% Pred. No. 1.7e+05;
Matches 3: Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 DFLT 7
DB 2 DFLT 7

RESULT 3
US-08-640-737-38
Sequence 38, Application US/08640737
Patent No. 6215044
GENERAL INFORMATION:
APPLICANT: ARROWSMITH, David A.
APPLICANT: HELLYER, Susan A.
APPLICANT: DE SILVA, Jacqueline
APPLICANT: WHITEMAN, Sally A.
TITLE OF INVENTION: Tomato xyloglucan Endo-Transglycosylase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/640,737
FILING DATE: 06-MAY-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/G94/02467
FILING DATE: 10-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9323225.4
FILING DATE: 10-NOV-1993
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: unknown

US-08-640-737-38

Query Match 54.3% Score 19: DB 4: Length 7;
Best Local Similarity 80.0% Pred. No. 1.7e+05;
Matches 4: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 DFLT 6
DB 3 DFLT 7

RESULT 4
US-08-405-230-10
Sequence 10, Application US/08405230
Patent No. 5707846
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: NISHIKAWA, Atsushi
APPLICANT: YAMAGUCHI, No. 5707846om1
TITLE OF INVENTION: NOVEL N-ACETYLGLUCOSAMINYL TRANSFERASE
TITLE OF INVENTION: GENE CODING THEREFOR AND PROCESS FOR PRODUCTION THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,230
FILING DATE: 16-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,736
FILING DATE: 23-AUG-1993
APPLICATION NUMBER: JP 4-245950
FILING DATE: 24-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-237118
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..5
OTHER INFORMATION:
OTHER INFORMATION: encoded by nucleotides 1-15 of SEQ ID NO. 7. *

Query Match 48.6% Score 17: DB 1: Length 5;
Best Local Similarity 100.0% Pred. No. 1.7e+05;
Matches 3: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-405-230-10

OY 1 TDF 3
111
Db 2 TDF 4

RESULT 5
US-08-910-990-10
; Sequence 10, Application US/08910990
; Patent No. 5834284
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: NISHIKAWA, Atsushi
APPLICANT: YAMAGUCHI, No. 5834284om1
TITLE OF INVENTION: NOVEL N-ACETYLGLUCOSAMINYL TRANSFERASE
TITLE OF INVENTION: GENE CODING THEREFOR AND PROCESS FOR PRODUCTION THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910.990
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,230
FILING DATE: 16-MAR-1995
APPLICATION NUMBER: US 08/110,736
FILING DATE: 23-AUG-1993
APPLICATION NUMBER: JP 4-245950
FILING DATE: 24-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-237118
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..5
OTHER INFORMATION: /note= "Amino acid sequence
encoded by nucleotides 1-15 of SEQ ID NO. 7."
US-08-910-990-10

Query Match 48.6%; Score 17; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDF 3
111
Db 2 TDF 4

RESULT 6
US-08-136-743B-29
; Sequence 29, Application US/08136743B
; Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
APPLICANT: Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inh1
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-136-743B-29

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDF 3
111
Db 5 TDF 7

RESULT 7
US-08-136-743B-30
; Sequence 30, Application US/08136743B
; Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
APPLICANT: Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide inh1
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-136-743B-30

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDF 3
|||
DB 5 TDF 7

RESULT 8
US-08-136-743B-31
Sequence 31, Application US/08136743B
Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhibit
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid

TOPOLOGY: linear
US-08-136-743B-31

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDF 3
|||
DB 5 TDF 7

RESULT 9
US-08-136-743B-32
Sequence 32, Application US/08136743B
Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-
TITLE OF INVENTION: cleotide Reductase, DNA Sequences Therefor and Peptide Inhi
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: The University of Pennsylvania
STREET: Suite 330
STREET: 3700 Market Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19104-3246
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/136,743B
FILING DATE: 10/14/93
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Monaco, Daniel A.
REGISTRATION NUMBER: 30,480
REFERENCE/DOCKET NUMBER: 3957-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-8383
TELEFAX: (215) 568-5549
TELEX: No. 5459063e
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-136-743B-32

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDF 3
|||
DB 5 TDF 7

RESULT 10
US-08-405-230-5
Sequence 5, Application US/08405230
Patent No. 5707846
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: NISHIKAWA, Atsushi

APPLICANT: YAMAGUCHI, NO. 5707846om1
TITLE OF INVENTION: NOVEL N-ACETYLGUCOSAMINYL TRANSFERASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/405,230
FILING DATE: 16-MAR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,736
FILING DATE: 23-AUG-1993
APPLICATION NUMBER: JP 4-245950
FILING DATE: 24-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-237118
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-405-230-5

Query Match 48.6%; Score 17; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDF 3
111
DB 2 TDF 4

RESULT 11
US-08-910-990-5
Sequence 5, Application US/08910990
Patent No. 5834284
GENERAL INFORMATION:
APPLICANT: TANIGUCHI, Naoyuki
APPLICANT: NISHIKAWA, Atsushi
APPLICANT: YAMAGUCHI, NO. 5834284om1
TITLE OF INVENTION: NOVEL N-ACETYLGUCOSAMINYL TRANSFERASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: George Mason Bldg., Washington & Prince Sts.
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,990
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,230
FILING DATE: 16-MAR-1995
APPLICATION NUMBER: US 08/110,736
FILING DATE: 23-AUG-1993
APPLICATION NUMBER: JP 4-245950
FILING DATE: 24-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-237118
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Crane-Feury, Sharon E
REGISTRATION NUMBER: 36,113
REFERENCE/DOCKET NUMBER: 001560-215
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-910-990-5

Query Match 48.6%; Score 17; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TDF 3
111
DB 2 TDF 4

RESULT 12
US-08-739-401A-6
Sequence 6, Application US/08739401A
Patent No. 5837461
GENERAL INFORMATION:
APPLICANT: Neitz, Maureen E.
APPLICANT: Neitz, John F.
TITLE OF INVENTION: DETECTION OF CONE-PHOTORECEPTOR-BASED
VISION DISORDERS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53202-4497
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/739,401A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Baker, Jean C.
REGISTRATION NUMBER: 35,433

REFERENCE/DOCKET NUMBER: 650053.91151
TELECOMMUNICATION INFORMATION:
TELEPHONE: (414) 277-5709
TELEFAX: (414) 271-3552
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-739-401A-6

Query Match 48.6%; Score 17; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PTLT 6
11:1
DB 2 FTVT 5

RESULT 13
PCT-US93-11703-72
Sequence 72, Application PC/TUS9311703
GENERAL INFORMATION:
APPLICANT: Chiron Mimotopes Pty. Ltd.
TITLE OF INVENTION: T-Cell Epitopes
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Grant D. Green
STREET: 4560 Horton St.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/11703
FILING DATE: 28-DEC-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/984,852
FILING DATE: 02-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Green, Grant D.
REGISTRATION NUMBER: 31,259
REFERENCE/DOCKET NUMBER: 0222.101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2706
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-11703-72

Query Match 48.6%; Score 17; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFT 4
111
DB 3 DFT 5

RESULT 14
US-08-180-209B-14
Sequence 14, Application US/08180209B
Patent No. 5593877
GENERAL INFORMATION:
APPLICANT: King, Te-Piao
TITLE OF INVENTION: CLONING AND RECOMBINANT PRODUCTION OF
TITLE OF INVENTION: VESPID VENOM ENZYMES, SUCH AS PHOSPHOLIPASE AND
TITLE OF INVENTION: HYALURONIDASE, AND IMMUNOLOGICAL THERAPIES BASED
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/180.209B
FILING DATE: 11-JAN-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/031,400
FILING DATE: 11-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-074 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-180-209B-14

Query Match 45.7%; Score 16; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFT 4
111
DB 2 TDLT 5

RESULT 15
US-08-385-745-14
Sequence 14, Application US/08385745
Patent No. 5612209
GENERAL INFORMATION:
APPLICANT: King, Te Piao

TITLE OF INVENTION: Cloning and Recombinant Production of
TITLE OF INVENTION: Vespid Venom Phospholipases, and Immunological Therapies
TITLE OF INVENTION: Based Thereon
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,745
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/031,400
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 3288-020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-385-745-14

Query Match 45.7%; Score 16; DB 1; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFT 4
DB 2 TDLT 5

Search completed: July 15, 2002, 13:26:14
Job time: 494 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:27:09 ; Search time 49.13 Seconds
(without alignments)
13.691 Million cell updates/sec

Title: US-09-712-819a-6

Perfect score: 33

Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues 455

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	45.5	7	2	S19630
2	13	39.4	7	2	A28709
3	12	36.4	5	2	T14910
4	11	33.3	5	2	PT0644
5	10	30.3	7	2	PS0254
6	9	27.3	3	3	T13892
7	9	27.3	4	2	I40804
8	9	27.3	4	2	T46627
9	9	27.3	5	2	A60521
10	9	27.3	5	2	E42364
11	9	27.3	5	2	A44955
12	9	27.3	5	2	S11127
13	9	27.3	5	2	PT0525
14	9	27.3	5	2	PT0577
15	9	27.3	5	2	PT0565
16	9	27.3	5	2	S69237
17	9	27.3	5	2	A60986
18	9	27.3	6	2	A43766
19	9	27.3	6	2	I37263
20	9	27.3	6	2	B26206
21	9	27.3	6	2	I65546
22	9	27.3	6	2	PT0518
23	9	27.3	6	2	PT0662
24	9	27.3	6	2	I49424
25	9	27.3	7	2	JN0859
26	9	27.3	7	2	A15398
27	9	27.3	7	2	B39127
28	9	27.3	7	2	S25266
29	9	27.3	7	2	S25266

30	9	27.3	7	2	A25269	sex pheromone CAN3
31	9	27.3	7	2	A30812	alpha-dextrin endo
32	9	27.3	7	2	PN0649	glucanase S-tran
33	9	27.3	7	2	S09066	major fat-globule
34	9	27.3	7	2	PN0150	hypothetical prote
35	9	27.3	7	2	S78024	hypothetical prote
36	9	27.3	7	2	E48394	ribosomal protein
37	9	27.3	7	2	I48086	glycoprotein compo
38	9	27.3	7	2	PT0671	DNA topoisomerase
39	9	27.3	7	2	S66442	T-cell receptor be
40	9	27.3	7	2	B48394	glutathione S-tran
41	8	24.2	4	2	I40505	major fat-globule
42	8	24.2	4	2	I39964	hypothetical prote
43	8	24.2	5	2	I39965	ribosomal protein
44	8	24.2	5	2	I39965	ribosomal protein
45	8	24.2	5	2	B22565	ribosomal protein

ALIGNMENTS

RESULT 1
S19630
ribosomal protein L30 - Streptomyces griseus (fragment)
C:Species: Streptomyces griseus
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 06-Jun-1993
C:Accession: S19630
R:Oehl, K.
Int. J. Syst. Bacteriol. 42, 144-150, 1992
A:Title: Electrophoretic heterogeneity of ribosomal protein AF-L30 among actinomycete
A:Reference number: S19630; MUID:92144363
A:Accession: S19630
A:Molecule type: protein
A:Residues: 1-7 <OCH>
A:Experimental source: strain IFO 13189
C:Superfamily: Escherichia coli ribosomal protein L30
C:Keywords: protein biosynthesis; ribosome

Query Match 45.5%; Score 15; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKISR 7
Db 3 LKITO 7
RESULT 2
A28709
phosphonacetaldehyde hydrolase - Bacillus cereus (fragment)
C:Species: Bacillus cereus
C>Date: 22-Aug-1988 #sequence_revision 22-Aug-1988 #text_change 30-Sep-1993
R:Olson, D.B.; Hepburn, T.W.; Moos, M.; Mariano, P.S.; Dunaway-Mariano, D.
Biochemistry 27, 2229-2234, 1988
A:Title: Investigation of the Bacillus cereus phosphonacetaldehyde hydrolase. Eviden
ide.
A:Reference number: A28709; MUID:88241058
A:Accession: A28709
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <OLS>

Query Match 39.4%; Score 13; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKI 5
Db 1 LKI 3

```
RESULT 3
T14910
hypothetical protein - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14910
R:Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A:Title: CPRF4, a novel plant bZIP protein of the CPRF family: comparative analysis of
A:Reference number: 218261; MUID:98265918
A:Accession: T14910
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5 <KR>
A:Cross-references: EMBL:Y10810; NID:93336904; PIDN:CAA71769.1; PID:93336905
A:Experimental source: ssp. Hamburger Schmitt

Query Match          36.4%; Score 12; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 ISR 7
    ||
Db 2 VSR 4

RESULT 4
PT0644
T-cell receptor beta chain V-D-J region (111-16) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0644
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A:Reference number: PT0509; MUID:91277601
A:Accession: PT0644
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <PEE>
A:Experimental source: newborn thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match          33.3%; Score 11; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FT 2
    ||
Db 4 FT 5

RESULT 5
PS0254
18K protein 5507 - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C:Accession: PS0254
R:Tangita, A.
submitted to JIPID, April 1993
A:Reference number: PS0206
A:Accession: PS0254
A:Molecule type: protein
A:Residues: 1-7 <TSU>
A:Experimental source: leaf, chloroplast, strain Nihonbare
A>Note: molecular weight 18K, pI 4.4

Query Match          30.3%; Score 10; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 LISR 7
    ||::
Db 1 LAIAK 5

RESULT 6
T13892
cytochrome-c oxidase (EC 1.9.3.1) chain I [imported] - river lamprey mitochondrion (f
C:Species: mitochondrion Lampetra fluviatilis (river lamprey)
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C:Accession: T13892
R:Delarbre, C.; Barriol, V.; Tillet, S.; Janvier, P.; Gachelin, G.
Mol. Biol. Evol. 14, 807-813, 1997
A:Title: The main features of the granulate mitochondrial DNA between the ND1 and the
A:Reference number: Z17775; MUID:97398704
A:Accession: T13892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3 <DEL>
A:Cross-references: EMBL:Y09528; NID:92340016; PIDN:CAA70721.1; PID:94379123
C:Genetics:
A:Genome: mitochondrion
A>Note: COI
C:Keywords: mitochondrion; oxidoreductase

Query Match          27.3%; Score 9; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TL 3
    ||
Db 2 TL 3

RESULT 7
I40804
endoglucanase F - Clostridium thermocellum (fragment)
C:Species: Clostridium thermocellum
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Aug-1996
C:Accession: I40804
R:Mishra, S.; Beguin, P.; Aubert, J.
J. Bacteriol. 173, 80-85, 1991
A:Title: Transcription of clostridium thermocellum endoglucanase genes celf and celd
A:Reference number: I40804; MUID:91100322
A:Accession: I40804
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4 <RES>
A:Cross-references: GB:M64363; NID:g144771
C:Genetics:
A:Gene: celf
A:Start codon: TTG

Query Match          27.3%; Score 9; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KI 5
    ||
Db 3 KI 4

RESULT 8
T46627
hypothetical protein c4 - loblolly pine
C:Species: Pinus taeda (loblolly pine)
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
C:Accession: T46627
R:Chang, S.; Puryear, J.; Funkhouser, E.A.; Newton, R.J.; Cairney, J.
```

submitted to the EMBL Data Library, July 1995
A:Description: Cloning of a chitinase homolog which lacks chitin binding sites and is dc
A:Reference number: 223105
A:Accession: U46627
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-4 <CH>
A:Cross-references: EMBL:U31309; NID:974285; PID:974292
A:Experimental source: strain 56PTxs6PT3; 8 month seedlings

Query Match 27.3%; Score 9; DB 2; Length 4;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 LKI 5
:||
Db 1 MKL 3

RESULT 9
A60521
glycogen phosphorylase (EC 2.4.1.1), muscle - mullet (Liza ramada) (fragment)
N:Alternate names: glycogen phosphorylase b
C:Species: Liza ramada
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-May-2000
C:Accession: A60521
R:Bonamusa, L.; Baanante, I.V.
Comp. Blochem. Physiol. B 95, 295-301, 1990
A:Title: Purification and characterization of glycogen phosphorylase B from skeletal mus
A:Reference number: A60521; MUID:90227907
A:Accession: A60521
A:Molecule type: protein
A:Residues: 1-5 <BON>
C:Superfamily: phosphorylase
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F3/Binding site: phosphate (Ser) (covalent) (by phosphorylase b kinase) #status experim

Query Match 27.3%; Score 9; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 KIS 6
:||
Db 1 QIS 3

RESULT 10
E42364
flagellar protein flir - Salmonella typhimurium (fragment)
C:Species: Salmonella typhimurium
C:Date: 24-Jul-1992 #sequence_revision 24-Jul-1992 #text_change 30-Sep-1993
C:Accession: E42364
R:Vogler, A.P.; Homma, M.; Irikura, V.M.; Macnab, R.M.
J. Bacteriol. 173, 3564-3572, 1991
A:Title: Salmonella typhimurium mutants defective in flagellar filament regrowth and seq
A:Reference number: A42364; MUID:91258342
A:Accession: E42364
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5 <VOG>
A:Cross-references: GB:M62408

Query Match 27.3%; Score 9; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
:||
Db 3 TL 4

RESULT 11
A44955
alkanal monooxygenase (PMN-linked) (EC 1.14.14.3) alpha chain - Vibrio harveyi (fragm
C:Species: Vibrio harveyi
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-May-2000
C:Accession: A44955
R:Paquette, O.; Yu, S.C.
Photochem. Photobiol. 50, 817-825, 1989
A:Title: Chemical modification and characterization of the alpha cysteine 106 at the
A:Reference number: A44955; MUID:90175700
A:Accession: A44955
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <PAQ>
C:Keywords: FMN; luminescence; monooxygenase; oxidoreductase

Query Match 27.3%; Score 9; DB 2; Length 5;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ISR 7
:||
Db 3 IXR 5

RESULT 12
S11127
phosphoprotein, bone - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 04-Mar-2000
C:Accession: S11127; S11128
R:Mikuni-Takagaki, Y.; Gilmer, M.J.
Biochem. J. 268, 585-591, 1990
A:Title: Post-translational processing of chicken bone phosphoproteins. Identificatio
A:Reference number: S11127; MUID:90303246
A:Accession: S11127
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-5 <MIK1>
A:Accession: S11128
A:Status: preliminary
A:Molecule type: protein
A:Residues: 'X', 2-5 <MIK2>
C:Keywords: phosphoprotein

Query Match 27.3%; Score 9; DB 2; Length 5;
Best Local Similarity 33.3%; Pred. No. 2.8e+05;
Matches 1; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 ISR 7
:||
Db 3 VSK 5

RESULT 13
PT0525
T-cell receptor beta chain V-D-J region (100-45) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C:Accession: PT0525
R:Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions
A:Reference number: PT0509; MUID:91277601
A:Accession: PT0525
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-5 <FEF>
A:Experimental source: adult thymus, strain BALB/c
C:Keywords: T-cell receptor

Query Match 27.3%: Score 9; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0;

OY 6 SR 7
 ||
 Db 2 SR 3

RESULT 14

PT0577
 T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0577; PT0574
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; MUID:91277601
 A:Accession: PT0577
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <FE2>
 A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC
 A:Accession: PT0574
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <FE2>
 A:Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q
 C:Keywords: T-cell receptor

Query Match 27.3%: Score 9; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0;

OY 6 SR 7
 ||
 Db 3 SR 4

RESULT 15

PT0565
 T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0565
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; MUID:91277601
 A:Accession: PT0565
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-5 <FE2>
 A:Experimental source: day 19 fetal thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 27.3%: Score 9; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0;

OY 6 SR 7
 ||
 Db 3 SR 4

Search completed: July 15, 2002, 13:27:09
 Job time: 439. sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:38:10 ; Search time 24.49 Seconds

(without alignments)
11.067 Million cell updates/sec

Title: US-09-712-819A-6

Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0

Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	33.3	5	1	RE31_LITRU
2	11	33.3	5	1	RE32_LITRU
3	10	30.3	7	1	GFRP_MOUSE
4	9	27.3	5	1	UF01_MOUSE
5	9	27.3	6	1	UN06_CLOPA
6	9	27.3	7	1	CCP1_ENTFA
7	9	27.3	7	1	CHOX_ALCSP
8	9	27.3	7	1	CIA_ENTFA
9	9	27.3	7	1	UC24_MAIZE
10	8	24.2	5	1	AL14_CARMA
11	8	24.2	5	1	PSK_DAUCA
12	8	24.2	7	1	AL13_CARMA
13	8	24.2	7	1	AL13_CARMA
14	8	24.2	7	1	AL14_CARMA
15	8	24.2	7	1	AL15_CARMA
16	8	24.2	7	1	AL17_CYPDO
17	8	24.2	7	1	CARP_MYED
18	8	24.2	7	1	FAR5_HIRME
19	7	21.2	3	1	LUXE_VIBFI
20	7	21.2	6	1	LOK1_LOCOMI
21	7	21.2	6	1	VP19_HSVIK
22	6	18.2	4	1	ACH1_ACHFU
23	6	18.2	4	1	FAR3_HIRME
24	6	18.2	4	1	FFK4_HIRME
25	6	18.2	4	1	FFK4_ANTEL
26	6	18.2	4	1	FLRF_HIRME
27	6	18.2	4	1	FLRF_ANTEL
28	6	18.2	4	1	FMRP_MACNT
29	6	18.2	4	1	FYRI_ANTEL
30	6	18.2	4	1	OCP1_OCTMI
31	6	18.2	5	1	EI03_LITRU
32	6	18.2	5	1	EI04_LITRU
33	6	18.2	5	1	FARP_ARTTR

34	6	18.2	5	1	PAP2_PARMA	P81864	parachirus
35	6	18.2	5	1	RE11_LITRU	P82070	litorea rub
36	6	18.2	5	1	RE21_LITRU	P82071	litorea rub
37	6	18.2	5	1	SUGA_ACHDO	P19991	acheta dome
38	6	18.2	5	1	TPIS_CANEA	P54714	canis fam11
39	6	18.2	5	1	UC22_MAIZE	P80628	zea mays (m
40	6	18.2	6	1	ACPH_RABIT	P25154	oryctolagus
41	6	18.2	6	1	CIP1_MYED	P13736	mytilus edu
42	6	18.2	6	1	CIP2_MYED	P13737	mytilus edu
43	6	18.2	6	1	EI01_LITRU	P82096	litorea rub
44	6	18.2	6	1	FARP_MONEX	P41966	monieza ex
45	6	18.2	7	1	EI05_LITRU	P82101	litorea rub

ALIGNMENTS

```
RESULT 1
AC RE31_LITRU STANDARD; PRT; 5 AA.
AC P82072;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidin 3.1.
OS Litorea rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litorea.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinboerner S.T., Wadnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litorea rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -!- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -!- MASS SPECTROMETRY: MM=655; METHOD=FAB.
KW Amphibian skin; Amidation.
PT MOD.RES 5 AA; 656 MM; 71A9C9CB10300000 CRC64;
SQ SEQUENCE
```

Query Match 33.3%; Score 11; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 FT 2
OY 11
Db 4 FT 5
```

RESULT 2
ID RE32_LITRU STANDARD; PRT; 5 AA.
AC P82073;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidin 3.2.
OS Litorea rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litorea.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;

RA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
 RT "Peptides from the skin glands of the Australian buzzing tree frog
 RT Litoria electrica. Comparison with the skin peptides from Litoria
 RT rubella.";
 RL Aust. J. Chem. 52:0-0(1999).
 CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
 CC ACTIVITY
 CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 KW Amphibian skin.
 SQ SEQUENCE 5 AA: 570 MW; 71A9C9C862A00000 CRC64;

Query Match 33.3%; Score 11; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 0;

OY 1 FT 2
 11
 DB 4 FT 5

RESULT 3
 GFRP_MOUSE STANDARD; PRT; 7 AA.
 ID GFRP_MOUSE
 AC P99025;
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE GTP cyclolysolase I feedback regulatory protein (P35) (Fragment).
 GN GCHFR OR GFRP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Van J.X.,
 RA Hoogland C., Appel R.D., Binz P.-A., Hochstrasser D.F.,
 RA Covthorne M.;
 RA Submitted (Aug-1998) to the SWISS-PROT data bank.
 CC -1- FUNCTION: MEDIATES TETRAHYDROBIOTERIN INHIBITION OF GTP
 CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
 CC (BY SIMILARITY).
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC SWISS-2DPAGE: P99025; MOUSE.
 DR INIT_MET 0
 FT NON_TER 7
 SQ SEQUENCE 7 AA: 806 MW; 71B5B057273B4700 CRC64;

Query Match 30.3%; Score 10; DB 1; Length 7;
 Best Local Similarity 75.0%; Pred. No. 1e+05; Indels 0; Gaps 0;
 Matches 3; Conservative 0; Mismatches 1;

OY 3 LKIS 6
 111
 DB 3 LKIS 6

RESULT 4
 UPO1_MOUSE STANDARD; PRT; 5 AA.
 ID UPO1_MOUSE
 AC P38639;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)
 DE Unknown protein from 2D-page of fibroblasts (P19) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE.
 RC TISSUE=Fibroblast;
 RX MEDLINE=95009907; PubMed=7523108;
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
 RT "Separation and sequencing of familial and novel murine proteins
 RT using preparative two-dimensional gel electrophoresis.";
 RL Electrophoresis 15:735-745(1994).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.6, ITS MW IS: 19 kDa.
 FT NON_TER 5
 SQ SEQUENCE 5 AA: 717 MW; 7364087043100000 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 5;
 Best Local Similarity 66.7%; Pred. No. 1e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 0; Mismatches 1;

OY 5 ISR 7
 11
 DB 2 ICR 4

RESULT 5
 UN06_CLOPA STANDARD; PRT; 6 AA.
 ID UN06_CLOPA
 AC P81351;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Unknown protein CP 6 from 2D-page (Fragment).
 OS Clostridium pasteurianum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1501;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=MS;
 RX MEDLINE=98291870; PubMed=9629918;
 RA Flengstrand R., Skjeldal L.;
 RT "Two-dimensional gel electrophoresis separation and N-terminal
 RT sequence analysis of proteins from Clostridium pasteurianum W5.";
 RL Electrophoresis 19:802-806(1998).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 5.0, ITS MW IS: 75.9 kDa.
 FT NON_TER 6
 SQ SEQUENCE 6 AA: 657 MW; 605B1DC1A45A8000 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 6;
 Best Local Similarity 50.0%; Pred. No. 1e+05; Indels 0; Gaps 0;
 Matches 2; Conservative 1; Mismatches 1;

OY 2 TLKI 5
 11
 DB 3 TAEI 6

RESULT 6
 CCF1_ENTFA STANDARD; PRT; 7 AA.
 ID CCF1_ENTFA
 AC P20104;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE Sex pheromone CCF10.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=89008313; PubMed=3139658;
 RA Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,

RA Asdit J.C., Dunny G.M., Suzuki A.;
 RT "Structure of cCF10, a peptide sex pheromone which induces
 RT conjugative transfer of the Streptococcus faecalis tetracycline
 RT resistance plasmid, pCF10.";
 RL J. Biol. Chem. 263:14574-14578(1988).
 CC -1- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
 CC HEMOLYSIN PLASMID PCF10.
 DR PIR: A30812; A30812.
 KW Pheromone.
 SO SEQUENCE 7 AA: 790 MW; 72C9D2C731B2C740 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TL 3
 DB 3 TL 4

RESULT 7
 CHOX_ALCSP STANDARD; PRT: 7 AA.
 ID P16101;
 AC 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DE Choline oxidase (EC 1.1.3.17) (Fragment).
 OS Alcaligenes sp.
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Alcaligenes.
 OX NCBI_TaxID=512;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=81006769; PubMed=6997283;
 RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
 RT "Identification and properties of the prosthetic group of choline
 RT oxidase from Alcaligenes sp.";
 RL J. Biochem. 88:197-203(1980).
 CC -1- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
 DR PIR: A15398; A15398.
 KW Oxidoreductase.
 FT NON_TER 7
 SO SEQUENCE 7 AA: 839 MW; 7415B1E457644A0 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SR 7
 DB 6 SR 7

RESULT 8
 CIA_ENTFA STANDARD; PRT: 7 AA.
 ID P11932;
 AC 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DE Sex pheromone CAM373 (Clumping-inducing agent) (CIA).
 OS Enterococcus faecalis (Streptococcus faecalis).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87005252; PubMed=3093276;
 RA Mori M., Tanaka H., Sakagami Y., Isogai A., Fujino M., Kitada C.,
 White B.A., An F.Y., Clewell D.B., Suzuki A.;

RT "Isolation and structure of the Streptococcus faecalis sex pheromone,
 RT cCM373.";
 RL FEBS Lett. 206:69-72(1986).
 CC -1- FUNCTION: CAM373 INDUCES MATING RESPONSE OF DONOR CELLS
 CC HARBORING PAM373.
 CC -1- MISCELLANEOUS: THE N-TERMINUS IS POSSIBLY RESPONSIBLE FOR
 CC SPECIFICITY OF PHEROMONES TO PLASMIDS.
 CC -1- SIMILARITY: C-TERMINAL TO THE PHEROMONES CPD1 AND CAD1.
 DR PIR: A25269; A25269.
 KW Pheromone.
 SO SEQUENCE 7 AA: 734 MW; 75BDD72059C05DB0 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
 DB 3 FTL 5

RESULT 9
 UC24_MAIZE STANDARD; PRT: 7 AA.
 ID P80630;
 AC 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DE 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 447)
 DE (Fragment).
 OS Zea mays (Maize).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RX TISSUE=Coleoptile;
 RA Penoulet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program.";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.0, ITS MW IS: 30.0 KDa.
 CC Maize-2DPAGE; P80630; COLEOPTILE.
 DR MaizeDB; 123956; -.
 KW NON_TER 1
 FT NON_TER 7
 SO SEQUENCE 7 AA: 665 MW; 6DC1B5B33DC1B5D0 CRC64;

Query Match 27.3%; Score 9; DB 1; Length 7;
 Best Local Similarity 66.7%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TLK 4
 DB 2 TAK 4

RESULT 10
 AL14_CARMA STANDARD; PRT: 5 AA.
 ID P81817;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Carcinustatin 14.
 OS Carcinus maenas (Common shore crab) (Green crab).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;

OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 5 5
SQ SEQUENCE 5 AA; 586 MW; 672879D5AB300000 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 5;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
DB 3 FGL 5

RESULT 11
PSK DAUCA
ID PSK DAUCA STANDARD; PRT: 5 AA.
AC p58261;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phytosulfokine-alpha (PSK-alpha) [Contains: Phytosulfokine-beta (PSK-beta)].
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OX Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE, AND IDENTIFICATION BY MASS SPECTROMETRY.
RC STRAIN=cv. US-Harumakigoun;
RX MEDLINE=20212743; PubMed=10750705;
RA Hanai H., Matsuno T., Yamamoto M., Matsubayashi Y., Kobayashi T.,
Kameda H., Sakagami Y.;
RT "A secreted peptide growth factor, phytosulfokine, acting as a
stimulatory factor of carrot somatic embryo formation.";
RL Plant Cell Physiol. 41:27-32(2000).
CC -1- FUNCTION: IN PRESENCE OF 2,4-D, STIMULATES PROLIFERATION OF THE
CELLS, BUT DOES NOT STIMULATE DIFFERENTIATION INTO THE SOMATIC
EMBRYOS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: SULFATION IS IMPORTANT FOR ACTIVITY AND FOR THE BINDING TO A
PUTATIVE MEMBRANE RECEPTOR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PHYTOSULFOKINE FAMILY.
KW Growth factor; Sulfation.
FT PEPTIDE 1 4
FT MOD_RES 1 1
FT MOD_RES 3 3
SQ SEQUENCE 5 AA; 687 MW; 76C1BB504B300000 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 5;
Best Local Similarity 50.0%; Pred. No. 1e+05;
Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTL 2
DB 3 FTL 4

RESULT 12
ALL2_CARMA
ID ALL2_CARMA STANDARD; PRT: 7 AA.
AC P81805;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas 2.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 7 7
SQ SEQUENCE 7 AA; 770 MW; 672879CDBC5DB870 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
DB 5 FGL 7

RESULT 13
ALL3_CARMA
ID ALL3_CARMA STANDARD; PRT: 7 AA.
AC P81806;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinus maenas 3.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA; 796 MW; 672879CDBC476B70 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
DB 5 FGL 7

Db 1 1
5 FGL 7

Search completed: July 15, 2002, 13:38:11
Job time: 711 sec

RESULT 14
ALL4_CARMA STANDARD; PRT; 7 AA.
AC P81807;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 4.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997)
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
SQ SEQUENCE 7 AA: 782 MW; 672879CDBC476AC0 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 1;

OY 1 FTL 3
1 1
Db 5 FGL 7

RESULT 15
ALL5_CARMA STANDARD; PRT; 7 AA.
ID ALL5_CARMA
AC P81808;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Carcinustatin 5.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidae; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997)
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Multigene family.
FT MOD_RES 7 AMIDATION.
SQ SEQUENCE 7 AA: 781 MW; 672879CDBC476420 CRC64;

Query Match 24.2%; Score 8; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1e+05; 1; Indels 0; Gaps 0;
Matches 2; Conservative 0; Mismatches 1;

OY 1 FTL 3

SWIAD
24/10
K

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd..

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:37:39 ; Search time 86.73 Seconds
(without alignments)
13.962 Million cell updates/sec

Title: US-09-712-819A-6
Perfect score: 33
Sequence: 1 FTIKISR 7

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 65

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13	39.4	6	10	P82181 spinacia ol
2	13	39.4	6	10	P82182 spinacia ol
3	13	33.3	5	13	P82072 litoria rub
4	11	33.3	5	13	P82073 litoria rub
5	11	33.3	7	8	O95945 saccharomye
6	10	30.3	6	10	P82541 spinacia ol
7	9	27.3	4	11	O08433 rattus norv
8	9	27.3	7	2	O07354 synecococc
9	9	27.3	7	4	O15897 rattus norv
10	9	27.3	7	10	P93233 lycopersico
11	9	27.3	7	11	O63480 rattus norv
12	9	27.3	7	11	O55184 rattus norv
13	8	24.2	7	2	O47029 enterobacte
14	8	24.2	7	2	P70804 azotobacter
15	8	24.2	7	12	O9Y1R0 human adeno
16	8	24.2	7	12	O9Y1Q9 human adeno

17	8	24.2	7	12	O9YVE3 human adeno
18	7	21.2	5	2	P83073 bacillus ce
19	7	21.2	7	2	P72081 nocardia la
20	7	21.2	7	12	O66205 transmissib
21	7	21.2	7	12	O9Y010 transmissib
22	6	18.2	5	13	P82070 litoria rub
23	6	18.2	5	13	P82071 litoria rub
24	6	18.2	5	13	P82072 litoria rub
25	6	18.2	5	13	P82099 litoria rub
26	6	18.2	5	13	P82100 litoria rub
27	6	18.2	6	13	P82096 litoria rub
28	6	18.2	7	2	O50556 actinobacill
29	6	18.2	7	2	O34028 sphingomona
30	6	18.2	7	2	O54248 streptomyce
31	6	18.2	7	4	O15903 homo sapien
32	6	18.2	7	6	O28742 oryctolagus
33	6	18.2	7	8	P92421 psathyrosta
34	6	18.2	7	8	P92385 hordeum mar
35	6	18.2	7	8	P92372 haynaldia v
36	6	18.2	7	8	P92403 lophopyrum
37	6	18.2	7	8	P92425 pseudoroegn
38	6	18.2	7	8	P92387 henaridia p
39	6	18.2	7	8	P92427 peridictyon
40	6	18.2	7	8	P92390 heteranthe
41	6	18.2	7	8	P92226 crithopsis
42	6	18.2	7	8	P92214 amblyopyrum
43	6	18.2	7	8	P92430 aegilops ta
44	6	18.2	7	8	P92221 bromus iner
45	6	18.2	7	8	P92442 taeniathearu
					P92381 hordeum bra

ALIGNMENTS

RESULT 1
ID P82181 PRELIMINARY: PRT: 6 AA.
AC P82181:
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 BETA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. ALVARO; TISSUE=LEAF;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Sudramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; Ribosomal_L10.
DR InterPro: IPR002363; Ribosomal_L10_eub.
DR Pfam: PF00466; Ribosomal_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER
SQ SEQUENCE 6 AA: 675 MW: 63218415B05DB000 CRC64;

Query Match 39.4%; Score 13; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 5 ISR 7
DB 2 ISR 4

RESULT 2
P82182 PRELIMINARY: PRT: 6 AA.
ID P82182
AC P82182;
DT 01-JUN-2000 (TREMBLrel. 14, Created)
DT 01-JUN-2000 (TREMBLrel. 14, Last sequence update)
DE CHLOROPLAST 50S RIBOSOMAL PROTEIN L10 GAMMA (FRAGMENT).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STRAIN=CV. ALVARO; TISSUE=LEAF;
MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR InterPro: IPR001790; RIBOSOMAL_L10.
DR InterPro: IPR02363; RIBOSOMAL_L10_eub.
DR Pfam: PF00466; RIBOSOMAL_L10; PARTIAL.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON_TER
SO SEQUENCE 6 AA; 675 MW; 63218415B05DB000 CRC64;

Query Match 39.4%; Score 13; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 ISR 7
DB 2 ISR 4

RESULT 3
P82072 PRELIMINARY: PRT: 5 AA.
ID P82072
AC P82072;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 3.1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=SKIN SECRETION;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowle J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litoria rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996)
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIOTIC ACTIVITY.

```

```

CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=655; METHOD=FAH.
KW Amphibian skin; Amidation.
FT MOD_RES 5
SO SEQUENCE 5 AA; 656 MW; 71A9C9CB10300000 CRC64;

Query Match 33.3%; Score 11; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
DB 4 FT 5

RESULT 4
P82073 PRELIMINARY: PRT: 5 AA.
ID P82073
AC P82073;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE RUBELIDIN 3.2.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OX Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN SECRETION;
RA Wabnitz P.A., Bowle J.H., Tyler M.J., Wallace J.C.;
RT "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella.";
RL Aust. J. Chem. 52:0-0(1999).
CC -1- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC ANTIOTIC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW Amphibian skin.
SO SEQUENCE 5 AA; 570 MW; 71A9C9CB62A00000 CRC64;

Query Match 33.3%; Score 11; DB 13; Length 5;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FT 2
DB 4 FT 5

RESULT 5
095945 PRELIMINARY: PRT: 7 AA.
ID 095945
AC 095945;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE INSIDE INTRON 5 (FRAGMENT).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RC MEDLINE=81069885; PubMed=6254986;
RX Bonitz S.G., Coruzzi G., Thalerfeld B., Tzagoloff A., Macino G.;
RA "Assembly of the mitochondrial membrane system: Structure and
RA nucleotide sequence of the gene coding for subunit 1 of yeast

```

RT cytochrome oxidase.";
 RL J. Biol. Chem. 255:11927-11941(1980).
 DR EMBL: V00694; CAA24066.1; -.
 KW Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 859 MW; 75B7232362CDC460 CRC64;

Query Match 33.3%; Score 11; DB 8; Length 7;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 KIS 6
 1:1
 DB 5 KLS 7

RESULT 6
 P82541 PRELIMINARY; PRT; 6 AA.
 AC P82541;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CHLOROPLAST 30S RIBOSOMAL PROTEIN S19 BETA (FRAGMENT).
 OS Spinacia oleracea (Spinach).
 OC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
 RC STRAIN=CV. ALMARO; TISSUE=LEAF;
 RX MEDLINE=20435797; PubMed=10874039;
 RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;
 RT "The plastid ribosomal proteins. Identification of all the proteins in
 RT the small subunit of an organelle ribosome (chloroplast).";
 RL J. Biol. Chem. 37:28455-28465(2000).
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
 CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
 CC -1- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
 CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN P.I. S19 BETA
 CC FORM IS THE MINOR BASIC FORM.
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.
 CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
 DR InterPro: IPR022222; Ribosomal_S19.
 DR Pfam: PF00203; Ribosomal_S19; PARTIAL.
 DR PRINTS: PRO0975; RIBOSOMAL_S19; PARTIAL.
 DR PROSITE: PS00323; RIBOSOMAL_S19; PARTIAL.
 KW Ribosomal protein; Chloroplast; rRNA-binding.
 FT NON_TER 6
 SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 30.3%; Score 10; DB 10; Length 6;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 TLK 4
 1:1
 DB 3 SLK 5

RESULT 7
 ID 008433 PRELIMINARY; PRT; 4 AA.
 AC 008433;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)

DE UDP-GLUCURONOSYLTRANSFERASE, MICROSOAL (EC 2.4.1.17) (UDPGT)
 DE (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GUNN;
 RX MEDLINE=91282758; PubMed=1840486;
 RA Sato H., Aono S., Kashiwamata S., Koizumi O.;

RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
 RT hyperbilirubinemic Gunn rat.";
 RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
 CC -1- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
 CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
 CC ENDOGENOUS COMPOUNDS.
 CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
 CC BETA-D-GLUCORONOSIDE.
 CC -1- SUBCELLULAR LOCATION: MICROSOAL.
 DR EMBL: S38636; AAB19259.1; -.
 KW Transferase; Glycosyltransferase; Microsome; Multigene family.
 FT NON_TER 1
 FT NON_TER 4
 SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 27.3%; Score 9; DB 11; Length 4;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LK 4
 1:1
 DB 3 LK 4

RESULT 8
 ID 007354 PRELIMINARY; PRT; 7 AA.
 AC 007354;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE NTRK (FRAGMENT).
 GN NTRK.
 OS Synechococcus sp. (strain PCC 8801 / Rf-1) (Cyanothecae PCC 8801).
 CC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
 OX NCBI_TaxID=41431;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rf-1;
 RX MEDLINE=99231861; PubMed=10217509;
 RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RT Rf-1.";
 RL Microbiology 145:743-753(1999).
 DR EMBL: AF003700; AAC35193.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 849 MW; 7412C72AA9D5B030 CRC64;

Query Match 27.3%; Score 9; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTL 3
 1:1
 DB 3 FDL 5

RESULT 9
 ID 015897

ID 015897 PRELIMINARY; PRT; 7 AA.
 AC 015897;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE (CLONE XPE6A11A) (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
 RA Coolbaugh M.I., Chinault C.A., Baldwin A., Lindsay E.A., Zhao Z.-Y.,
 RA Caskey C.T.H.;
 RT "Isolation of chromosome-specific genes by reciprocal probing of
 RT arrayed cDNAs and cosmid libraries";
 RL Hum. Mol. Genet. 0:0-0(1995).
 DR EMBL: L32077; AAA73887.1; -.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 814 MW; 672B1D3372046B0 CRC64;

Query Match 27.3%; Score 9; DB 4; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 LK 4
 ||
 Db 3 LK 4

RESULT 10
 P93233 PRELIMINARY; PRT; 7 AA.
 ID P93233;
 AC P93233;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE (EC 4.4.1.14)
 DE (FRAGMENT).
 GN LE-ACSIB.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eusterids I; Solanales; Solanaceae; Solanum.
 NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97351561; PubMed=9707843;
 RA Oetiker J.H., Olson D.C., Shiu O.Y., Yang S.F.;
 RT "Differential induction of seven 1-aminocyclopropane-1-carboxylate
 RT synthase genes by elicitor in suspension cultures of tomato
 RT (Lycopersicon esculentum).";
 RL Plant Mol. Biol. 34:275-286(1997).
 DR EMBL: U75692; AAC49682.1; -.
 KW Lyase.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 828 MW; 71B412C7377415D0 CRC64;

Query Match 27.3%; Score 9; DB 10; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 SR 7
 ||
 Db 1, SR 2

RESULT 11 ;

063480
 ID 063480 PRELIMINARY; PRT; 7 AA.
 AC 063480;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TR4-NS ORPHAN RECEPTOR (FRAGMENT).
 GN TR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96198747; PubMed=8612486;
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
 RA Detera-Wadleigh S.D.;
 RT "Splice variants of rat TR4 orphan receptor: differential expression
 RT of novel sequences in the 5'-untranslated region and C-terminal
 RT domain.";
 RL Endocrinology 137:1562-1571(1996).
 DR EMBL: U59125; AAB02827.1; -.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 758 MW; 672A87864005350 CRC64;

Query Match 27.3%; Score 9; DB 11; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KI 5
 ||
 Db 1 KI 2

RESULT 12
 055184 PRELIMINARY; PRT; 7 AA.
 ID 055184;
 AC 055184;
 DT 01-JUN-1998 (TREMBlrel. 06, Created)
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE ORPHAN RECEPTOR TR4-NS (FRAGMENT).
 GN TR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=SPRAGUE-DAWLEY;
 RC MEDLINE=96198747; PubMed=8612486;
 RA Yoshikawa T., Makino S., Gao X.M., Xing G.Q., Chuang D.M.,
 RA Detera-Wadleigh S.D.;
 RT "Splice variants of rat TR4 orphan receptor: differential expression
 RT of novel sequences in the 5'-untranslated region and C-terminal
 RT domain.";
 RL Endocrinology 137:1562-1571(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=96299786; PubMed=8661150;
 RA Yoshikawa T., DuPont B.R., Leach R.J., Detera-Wadleigh S.D.;
 RT "New variants of the human and rat nuclear hormone receptor, TR4:
 RT expression and chromosomal localization of the human gene.";
 RL Genomics 35:361-366(1996).
 DR EMBL: U59454; AAB91433.1; -.
 KW Receptor.
 FT NON_TER 1 1
 FT NON_TER 7 7
 SQ SEQUENCE 7 AA; 663 MW; 6DDA8787E805350 CRC64;

Query Match 27.3%; Score 9; DB 11; Length 7;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KI 5
 1:
 1:
 Db 1 KI 2

RESULT 13
 Q47029 PRELIMINARY; PRT; 7 AA.
 ID Q47029;
 AC Q47029;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE ADP A1 PROTEIN (FRAGMENT).
 GN ADP A1.
 OS Enterobacter cloacae.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Enterobacter.
 OX NCBI_TaxID=550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94079349; PubMed=8257126;
 RA Rather P.N., Mann P.A., Mierzwa R., Hare R.S., Miller G.H., Shaw K.J.;
 RT "Analysis of the aac(3)-Vla gene encoding a novel 3-N-
 RT acetyltransferase."
 RL Antimicrob. Agents Chemother. 37:2074-2079(1993).
 DR EMBL: M88012; AAA16193.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 24.2%; Score 8; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 KI 5
 1:
 1:
 Db 3 KV 4

RESULT 14
 P70804 PRELIMINARY; PRT; 7 AA.
 ID P70804;
 AC P70804;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ALGT PROTEIN (FRAGMENT).
 GN ALGT.
 OS Azotobacter vinelandii.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Azotobacter.
 OX NCBI_TaxID=354;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=E;
 RX MEDLINE=96427318; PubMed=8830682;
 RA Rehm B.H.A., Ertesvag H., Valla S.;
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algC) is
 RT part of an alg gene cluster physically organized in a manner similar
 RT to that in Pseudomonas aeruginosa."
 RL J. Bacteriol. 178:5884-5889(1996).
 DR EMBL: X87973; CAA61230.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 24.2%; Score 8; DB 2; Length 7;
 Best Local Similarity 40.0%; Pred. No. 5.6e+05;
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TKIS 6
 1:
 1:
 Db 2 TVSSS 6

RESULT 15
 Q9YIR0 PRELIMINARY; PRT; 7 AA.
 ID Q9YIR0;
 AC Q9YIR0;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE PVI CORE PROTEIN (FRAGMENT).
 GN PVI.
 OS Human adenovirus type 7a.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=85755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KN T96-0620, S-1058, AND CL 68578;
 RA Crawford-Miksza L.R., Nang R.N., Schnurr D.P.;
 RT "Molecular surveillance of strain variation in adenoviruses causing
 RT acute respiratory disease, AV 4 and AV 7a."
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF065068; AAD03668.1; -.
 DR EMBL: AF065066; AAD03664.1; -.
 DR EMBL: AF065067; AAD03666.1; -.
 FT NON_TER 1
 SQ SEQUENCE 7 AA; 980 MW; 7B5EA414140322A0 CRC64;

Query Match 24.2%; Score 8; DB 12; Length 7;
 Best Local Similarity 33.3%; Pred. No. 5.6e+05;
 Matches 1; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 5 ISR 7
 1:
 1:
 Db 1 VKR 3

Search completed: July 15, 2002, 13:37:39
 Job time: 729 sec

10

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:25:27 ; Search time 107.71 Seconds

(without alignments)
7.219 Million cell updates/sec

Title: US-09-712-819A-6

Perfect score: 33

Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 52936

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq 032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	69.7	7	20	AAV40736
2	23	69.7	7	21	AAV30074
3	22	66.7	7	20	AAV40738
4	22	66.7	7	21	AAV30076
5	21	63.6	7	20	AAV42013
6	20	60.6	7	16	AAV81848
7	20	60.6	7	20	AAV41889
8	20	60.6	7	22	AAV55870
9	20	60.6	7	22	AAV55981
10	20	60.6	7	22	AAV56283
11	20	60.6	7	22	AAV52190

12	20	60.6	7	22	AAV52355	Human APT-125 trypt
13	20	60.6	7	22	AAU28602	DPI tryptic digest
14	20	60.6	7	22	AAU24969	Schizophrenia-Asso
15	20	60.6	7	22	AAU26249	Depression-Asso
16	20	60.6	7	22	AAU15313	Schizophrenia-Asso
17	19	57.6	7	20	AAV40737	S4 derivative #11,
18	19	57.6	7	21	AAV30075	Scaffold protein S
19	17	51.5	6	20	AAV84431	HIV-1 nucleic acid
20	17	51.5	7	11	AAV07656	Ribonuclease reduct
21	17	51.5	7	17	AAV97970	Antigenic fragment
22	17	51.5	7	19	AAV69269	Haemagglutinin hea
23	16	48.5	5	16	AAV75584	gp120 binding Fab
24	16	48.5	6	10	AAV93345	Portion of myc var
25	16	48.5	6	18	AAV39453	Human T cell epito
26	16	48.5	6	20	AAV49184	TATA box recognizi
27	16	48.5	6	21	AAV38999	Human secreted pep
28	16	48.5	6	22	AAV63074	A Hepatitis A viru
29	16	48.5	7	16	AAV72775	Mammalian ribonuc
30	16	48.5	7	20	AAV82668	Cauliflower cludase
31	16	48.5	7	21	AAV02931	Nucleotide-binding
32	16	48.5	7	21	AAV83858	Ribonucleotide red
33	16	48.5	7	22	AAV10492	Humanised Ab clone
34	15	45.5	4	19	AAV50120	Pan DR binding pep
35	15	45.5	4	21	AAV82816	Cathepsin-B inhibi
36	15	45.5	5	5	AAV40546	Sequence of peptid
37	15	45.5	5	8	AAV71651	Diuretic peptide d
38	15	45.5	5	16	AAV75578	gp120 binding Fab
39	15	45.5	5	16	AAV75587	gp120 binding Fab
40	15	45.5	5	16	AAV72783	Mammalian ribonuc
41	15	45.5	5	21	AAV83890	Mammalian ribonuc
42	15	45.5	5	22	AAV62805	Beta-amyloid pepti
43	15	45.5	6	11	AAV06664	Retroviral proteas
44	15	45.5	6	16	AAV72782	Mammalian ribonuc
45	15	45.5	6	16	AAV82924	Non-RGD, non-ITSGR

ALIGNMENTS

RESULT 1	
AAV40736	standard; peptide: 7 AA.
ID	AAV40736
XX	
AC	AAV40736;
XX	
DT	01-DEC-1999 (first entry)
XX	
DE	S4 derivative #10, beta strand of scaffold protein structure.
XX	
KW	Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW	tumour; chemotherapeutic agent.
OS	Synthetic.
XX	
PN	EP947582-AI.
XX	
PD	06-OCT-1999.
XX	
PF	31-MAR-1996; 98EP-0870065.
XX	
PR	31-MAR-1996; 98EP-0870065.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX	
DR	WPI, 1999-542958/46.
XX	
PT	New scaffold protein, useful for stabilizing antigens used as vaccines
XX	
PS	
XX	Disclosure; Page 6; 105pp; English.
XX	

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10⁸ alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX Sequence 7 AA:

Query Match 69.7%; Score 23; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKIS 6
| | | | |
Db 2 ftlsls 7

RESULT 2

AAB30074
ID AAB30074 standard; Peptide; 7 AA.

AC AAB30074;

DT 09-FEB-2001 (first entry)

XX Scaffold protein SCA S4 peptide SEQ ID NO: 135.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

XX Synthetic.

PN WO200060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

XX Scaffold composed of single-chain polypeptide having beta sandwich
XX architecture carrying new and randomized peptide sequences useful as
XX supporting framework and carrying antigen- or receptor binding
XX fragments -

PS Disclosure; Page 15; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB2930-B2939 were used in the
CC production of the proteins of the invention.

XX Sequence 7 AA:

Query Match 69.7%; Score 23; DB 21; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKIS 6
| | | | |
Db 2 ftlsls 7

RESULT 3

AAY40738
ID AAY40738 standard; peptide; 7 AA.

AC AAY40738;

DT 01-DEC-1999 (first entry)

DE S4 derivative #12, beta strand of scaffold protein structure.

KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.

XX Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

XX New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure; Page 6; 105pp; English.

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10⁸ alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will

CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX
SQ Sequence 7 AA;

Query Match 66.7%; Score 22; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTKIS 6
| | | | |
Db 2 ftkis 7

RESULT 4

AAB30076
ID AAB30076 standard; Peptide: 7 AA.

XX
AC AAB30076;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA s4 peptide SEQ ID NO: 137.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

XX
OS Synthetic.

PN WO00060070-A1.

PD 12-OCT-2000.

XX
PF 01-APR-1999; 99WO-EP02283.

XX
PR 01-APR-1999; 99WO-EP02283.

XX
PA (INNO-) INNOGENETICS NV.

XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

XX
DR WPI: 2000-665002/64.

XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments -

XX
PS Disclosure: Page 15; 68pp; English.

XX
CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.

XX
SQ Sequence 7 AA;

Query Match 66.7%; Score 22; DB 21; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTKIS 6
| | | | |

Db 2 ftkis 7

RESULT 5

AA42013
ID AA42013 standard; Peptide: 7 AA.

XX
AC AA42013;

DT 09-DEC-1999 (first entry)

DE Rheumatoid arthritis diagnostic protein isoform peptide #164.

KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
KW rheumatoid arthritis diagnostic feature; RPI; synovial fluid;

KW rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.

XX
OS Homo sapiens.

PN WO947925-A2.

PD 23-SEP-1999.

XX
PF 15-MAR-1999; 99WO-GB00763.

XX
PR 13-MAR-1998; 98GB-0005477.

XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX
PI Parekh RB, Patel TP, Townsend RR;

XX
DR WPI: 1999-571871/48.

XX
PT Diagnosis of human rheumatoid arthritis by two-dimensional
PT electrophoresis -

XX
PS Disclosure: Page 21; 157pp; English.

XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a
CC two-dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC compounds that promote or inhibit their activity, which are then used as
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AA42103 represent expression reference protein isoform peptides and
CC AA425066 to AA425068 represent degenerate probes for RPIs, which are all
CC used in the exemplification of the present invention.

XX
SQ Sequence 7 AA;

Query Match 63.6%; Score 21; DB 20; Length 7;
Best Local Similarity 83.3%; Pred. No. 6.4e+05;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 TIKIS 7
| | | | |
Db 2 tkisr 7

```

RESULT 6
AAR81848
ID AAR81848 standard; peptide: 7 AA.
XX
XX AAR81848;
XX
XX
XX 16-MAY-1996 (first entry)
XX
XX Human afamin tryptic fragment FX20.
XX
XX
XX Human: afamin; serum protein family; albumin; alpha-fetoprotein; plasma;
XX vitamin D binding protein; homology; post-translational processing;
XX chromatography; Primer: PCR; amplification; probe: rheumatoid arthritis;
XX ischaemia-reperfusion injury; ARDS; cardiopulmonary bypass; sepsis;
XX toxic plasma substance; inflammation.
XX
XX Homo sapiens.
XX
XX WO9527059-A1.
XX
XX 12-OCT-1995.
XX
XX 31-MAR-1995; 95WO-US04075.
XX
XX 31-MAR-1994; 94US-0222619.
XX
XX (AMGE-) AMGEN INC.
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Lichenstein HS, Lyons DE, Wright SD, Wurfel MM;
XX WPI; 1995-358634/46.
XX
XX Human afamin or a variant and poly-nucleotide(s) encoding it - a
XX human serum protein with activities in common with other members of
XX this family.
XX
XX Example 3; Page 45; 97pp; English.
XX
XX Peptides AAR81847-54 are tryptic peptide fragments from human afamin
XX (AAR81845) novel member of the human serum protein family. The
XX fragments were used to design primers and probes (AAT00786-98) for the
XX cloning of the vitamin gene (AAT00785) from human liver cDNA. Afamin is
XX thought to have similar properties to human albumin, alpha-fetoprotein
XX and vitamin D binding protein due to homology with these proteins. The
XX gene encodes a mature protein of 66576 daltons without post-translational
XX processing (ca. 87000 daltons with post-translational processing). The
XX protein was isolated from human plasma by a conventional chromatographic
XX methods. The protein can be used to ameliorate ischaemia-reperfusion
XX injury, rheumatoid arthritis, ARDS, cardiopulmonary bypass, sepsis, toxic
XX plasma substances released after inflammation, etc.
XX
XX Sequence 7 AA:
SQ

```

Query Match 60.6%; Score 20; DB 16; Length 7;
 Best Local Similarity 57.1%; Pred. No. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 FTLKISR 7
   | : | |
DB 1 lfefysr 7

```

```

RESULT 7
AAV41889
ID AAV41889 standard; Peptide: 7 AA.
XX
XX AAV41889;
XX
XX
XX 09-DEC-1999 (first entry)
DT

```

```

XX
XX Rheumatoid arthritis diagnostic protein isoform peptide #40.
DE
XX
XX Human: rheumatoid arthritis; RA; diagnosis; RPI; RADF; detection;
XX rheumatoid arthritis diagnostic feature; ERF; synovial fluid;
XX rheumatoid arthritis diagnostic protein isoform; screening;
XX expression reference protein isoform; prognosis.
XX
XX Homo sapiens.
XX
XX WO9447925-A2.
XX
XX 23-SEP-1999.
XX
XX 15-MAR-1999; 99WO-GB00763.
XX
XX 13-MAR-1998; 98GB-0005477.
XX
XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
XX Parekh RB, Patel TP, Townsend RR;
XX WPI; 1999-571871/48.
XX
XX Diagnosis of human rheumatoid arthritis by two-dimensional
XX electrophoresis -
XX
XX Disclosure; Page 18; 157pp; English.
XX
XX A method has been developed for the diagnosis of human rheumatoid
XX arthritis (RA) using two-dimensional electrophoresis to generate a
XX two-dimensional array of features. The method can be used for screening,
XX diagnosis and prognosis of RA in a subject or for monitoring the effect
XX of an anti-RA drug or therapy administered to a subject. The method
XX comprises: (a) analysing a sample of serum or plasma and optionally
XX synovial fluid by two-dimensional electrophoresis, to generate a two-
XX dimensional array of features; (b) identifying at least one chosen
XX feature whose relative abundance correlates with the presence or absence
XX of RA; and (c) comparing the abundance of each chosen feature in the
XX sample with the abundance of that chosen feature in serum or plasma from
XX one or more persons without RA, where the relative abundance of the
XX chosen feature or features in the sample indicates the presence or
XX absence of RA in the subject. The method can also be used in clinical
XX studies for testing drugs for therapy of RA, for purification of RA-
XX diagnostic protein isoforms (RPIs), and for production of antibodies to
XX RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
XX compounds that promote or inhibit their activity, which are then used as
XX RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
XX protocols. AAV41844 to AAV42100 represent RPI peptides, AAV42101 to
XX CC AAV42103 represent expression reference protein isoform peptides and
XX CC AA225066 to AA225068 represent degenerate probes for RPIs, which are all
XX used in the exemplification of the present invention.
XX
XX Sequence 7 AA:
SQ

```

Query Match 60.6%; Score 20; DB 20; Length 7;
 Best Local Similarity 42.9%; Pred. No. 6.4e+05;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 FTLKISR 7
   | : | |
DB 1 ycfelsr 7

```

```

RESULT 8
ABB55870
ID ABB55870 standard; Peptide: 7 AA.
XX
XX ABB55870;
XX
XX
XX 15-FEB-2002 (first entry)
DT
XX

```

DE Vascular dementia-associated protein isoform (VPI) 70.
XX
XX Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KM diagnosis; prognosis; gene therapy.
XX
XX Homo sapiens.
OS
PN MO200169261-A2.
FN
PD 20-SEP-2001.
PP
PF 14-MAR-2001; 2001WO-GB01106.
PR 15-MAR-2000; 2000GB-0006285.
PR 24-NOV-2000; 2000GB-0028734.
PR 28-NOV-2000; 2000US-0724391.
XA
PA (OXFO-) OXFORD GLYCOSCIENTICES UK LTD.
XX
XX Herath HMAC, Parekh RB, Rohlf C;
PI WPI; 2001-557937/62.
DR
XX Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy, for
PT comprises analysing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD -
XX
PS Claim 6; Page 31; 151pp: English.

CC The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for
CC monitoring the effect of therapy administered to a subject having VD.
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
CC useful for the treatment of VD and for gene therapy.
CX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 42.9%; Pred. No. 6,4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 FTUKISR 7
I I::II
DB 1 ycfelr 7

RESULT 9
ABB55981
ID ABB55981 standard; Peptide; 7 AA.
XX
XX AC ABB55981;
XX
DT 15-FEB-2002 (first entry)
DE Vascular dementia-associated protein isoform (VPI) 181.
XX
XX Vascular Dementia; VD; VD-associated protein isoform; VPI; screening;
KM diagnosis; prognosis; gene therapy.
XX
OS Homo sapiens.
PN WO200169261-A2.

[illegible]

```

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA
XX
PI Herath HMAc, Parekh RB, Rohlf C;
XX
DR WPI; 2001-557937/62.
XX
PT Screening, diagnosis or prognosis of vascular dementia (VD), useful for
PT determining stage of VD and monitoring the effect of VD therapy;
PT comprises analysing body fluid by 2-dimensional electrophoresis for
PT features correlated with VD -
XX
PS Claim 6; Page 40; 151pp; English.
XX
XX The invention relates to screening, diagnosis or prognosis of Vascular
CC Dementia (VD) in a subject comprising analysing body fluid from the
CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
CC features containing at least one chosen feature whose relative abundance
CC correlates with the presence, absence, stage or severity of VD or
CC predicts the onset or course of VD, especially detecting in a sample of
CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
CC specification. Detecting VD-associated features and VPI is useful for the
CC screening, diagnosis or prognosis of VD, for determining the stage or
CC severity of VD, for identifying a subject at risk of VD or for
CC monitoring the effect of therapy administered to a subject having VD.
CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
CC useful for the treatment of VD and for gene therapy.
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 42.9%; Pred. No. 6.4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTLKISR 7
   | : | | |
DB 1 ytfelst 7

RESULT 11
ABB52190
ID ABB52190 standard; Peptide: 7 AA.
XX
AC ABB52190;
XX
DT 08-FEB-2002 (first entry)
XX
DE Human API-146 tryptic digest peptide #1.
XX
XX Human; neuroprotective; nootropic; gene therapy; vaccine;
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS Homo sapiens.
XX
PN WO200175454-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US10908.
XX
PR 03-APR-2000; 2000US-194504P.
XX
PR 28-NOV-2000; 2000US-253647P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA (PFI2 ) PFIZER INC.
XX
PI Durham KL, Friedman DL, Herath HMAc, Kimmel LH, Parekh RB;
PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;
PI Townsend RR, White F, Williams SA.

```

```

XX WPI; 2001-639384/73.
XX
DR Screening for Alzheimer's disease in a mammal, by making
XX two-dimensional array of a feature whose relative abundance correlates
XX with disease, and comparing with abundance of the feature in samples of
XX healthy persons -
XX
PS Example; Page 30; 162pp; English.
XX
XX The invention relates to methods for the screening, diagnosis and
CC prognosis of Alzheimer's disease. The methods involve the detection
CC of Alzheimer's Disease-Associated Features (AFs) and Alzheimer's
CC Disease-Associated Protein Isoforms (APIs) in cerebrospinal fluid,
CC serum or plasma. The abundance of the AFs and APIs is then
CC normalised to an Expression Reference Protein Isoform (ERPI) in
CC order to determine whether a patient is suffering from, or has
CC a predisposition to, Alzheimer's disease. The relative abundance of
CC the AFs and APIs correlates with the severity of Alzheimer's disease.
CC The present sequence is a peptide produced from an API by proteolysis.
XX
SQ Sequence 7 AA;

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 57.1%; Pred. No. 6.4e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 FTLKISR 7
   | : | | |
DB 1 ttfeyst 7

RESULT 12
ABB52355
ID ABB52355 standard; Peptide: 7 AA.
XX
AC ABB52355;
XX
DT 08-FEB-2002 (first entry)
XX
DE Human API-125 tryptic digest peptide #8.
XX
XX Human; neuroprotective; nootropic; gene therapy; vaccine;
KW Alzheimer's disease; Alzheimer's Disease-Associated Feature; AF;
KW Alzheimer's Disease-Associated Protein Isoform; API; tryptic digest;
KW Expression Reference Protein Isoform; ERPI; proteolysis.
XX
OS Homo sapiens.
XX
PN WO200175454-A2.
XX
PD 11-OCT-2001.
XX
PF 03-APR-2001; 2001WO-US10908.
XX
PR 03-APR-2000; 2000US-194504P.
XX
PR 28-NOV-2000; 2000US-253647P.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
PA (PFI2 ) PFIZER INC.
XX
PI Durham KL, Friedman DL, Herath HMAc, Kimmel LH, Parekh RB;
PI Potter DM, Rohlf C, Silber BM, Stiger TR, Sunderland PT;
PI Townsend RR, White F, Williams SA;
XX
DR WPI; 2001-639384/73.
XX
XX Screening for Alzheimer's disease in a mammal, by making
XX two-dimensional array of a feature whose relative abundance correlates
XX with disease, and comparing with abundance of the feature in samples of
XX healthy persons -
XX

```


PS Example; Page 34; 162pp; English.
XX
CC The invention relates to methods for the screening, diagnosis and
CC prognosis of Alzheimer's disease. The methods involve the detection
CC of Alzheimer's Disease-Associated Features (AFs) and Alzheimer's
CC Disease-Associated Protein Isoforms (APIs) in cerebrospinal fluid,
CC serum or plasma. The abundance of the AFs and APIs is then
CC normalised to an Expression Reference Protein Isoform (ERPI) in
CC order to determine whether a patient is suffering from, or has
CC a predisposition to, Alzheimer's Disease. The relative abundance of
CC the AFs and APIs correlates with the severity of Alzheimer's Disease.
CC The present sequence is a peptide produced from an API by proteolysis.
XX
SQ Sequence 7 AA:

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 42.9%; Pred. No. 6.4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKISR 7
:|::||
Db 1 ytfelsr 7

RESULT 13
AAU28602
ID AAU28602 standard; Peptide; 7 AA.
XX
AC AAU28602;
XX
DT 03-JAN-2002 (first entry)
XX
DE DPI tryptic digest peptide #199.
XX
KM Human: depression associated protein isoform; tryptic digest peptide;
KM DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
KM neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
KM maniac-depressive illness; schizoaffective disorder.
XX
OS Homo sapiens.
XX
PN WO200162787-A1.
XX
PD 30-AUG-2001.
XX
PE 23-FEB-2001; 2001WO-GB00786.
XX
PR 24-FEB-2000; 2000GB-0004412.
PR 08-DEC-2000; 2000GB-0030050.
PR 12-DEC-2000; 2000US-0254830.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570626/64.
XX
PT Novel nucleic acid encoding a protein associated with bipolar affective
PT disorder, which is used for diagnosis, prophylaxis and therapy of
PT neuropsychiatric disorders, such as bipolar affective disorder -
XX
PS Disclosure; Page 34; 153pp; English.
XX
CC The present invention relates to the identification of depression
CC associated protein isoforms (DPis), particularly the tryptic digest
CC peptides of these proteins. Some of the DPis (AAU28404-AAU28625)
CC described are decreased in the cerebrospinal fluid (CSF) of BAD
CC (bipolar affective disorder) subjects, whilst other DPis
CC (AAU28626-AAU28887) are increased in BAD subjects. Also described
CC are peptide sequences identified from DPI-45 and DPI-213 and the
CC nucleic acid sequence they are encoded by. The sequences of the
CC invention are useful for clinical screening, diagnosis, prognosis,

CC therapy and prophylaxis of neuropsychiatric disorders e.g. BAD (also
CC known as bipolar mood disorder, BP), maniac-depressive illnesses,
CC attention deficit disorders, schizoaffective disorders, and unipolar
CC affective disorders. The present sequence represents one of the DPI
CC tryptic digest peptides of the present invention.
XX
SQ Sequence 7 AA:

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 42.9%; Pred. No. 6.4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKISR 7
:|::||
Db 1 ytfelsr 7

RESULT 14
AAU24969
ID AAU24969 standard; Peptide; 7 AA.
XX
AC AAU24969;
XX
DT 18-DEC-2001 (first entry)
XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #198.
XX
KM Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-240;
KM neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS Homo sapiens.
XX
PN WO200162785-A2.
XX
PD 30-AUG-2001.
XX
PE 23-FEB-2001; 2001WO-GB00792.
XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMC, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570624/64.
XX
PT New schizophrenia associated protein isoforms and encoding nucleic acid
PT molecules, useful for treatment, diagnosis and prognosis of
PT schizophrenia and screening for potential drugs for treatment and new
PT drug targets -
XX
PS Disclosure; Page 32; 148pp; English.
XX
CC The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs.
XX
SQ Sequence 7 AA:

Query Match 60.6%; Score 20; DB 22; Length 7;
Best Local Similarity 42.9%; Pred. No. 6.4e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKISR 7
:|::||

Db 1 ytfelst 7

Search completed: July 15, 2002, 13:25:27
Job time: 1458 sec

RESULT 15

ID AAU26249 standard; Peptide: 7 AA.

AC AAU26249;

DT 18-DEC-2001 (first entry)

DE Depression-Associated Protein Isoform DPI-208.

KW Human; Bipolar Affective Disorder; BAD; Depression-Associated feature;
 KW DF; Depression-Associated Protein Isoform; DPI; Cerebro-spinal fluid;
 KW CSF; antidepressant; antimanic; nootropic; tranquilizer; neuroleptic;
 KW attention deficient disorder; schizoaffective disorder;
 KW unipolar affective disorder.

XX Homo sapiens.

OS WO200163294-A2.

PN 30-AUG-2001.

XX 23-FEB-2001; 2001WO-GB00791.

XX 24-FEB-2000; 2000GB-0004412.

PR 08-DEC-2000; 2000GB-0030050.

PR 12-DEC-2000; 2000US-0254830.

XX (OXFO-) OXFORD GLYCOSCIENCES UK LTD.

XX Herath HMC, Parekh RB, Rohlf C;

DR WPI: 2001-582081/65.

PT Preparation for diagnosing or treating bipolar affected disorder (BAD)
 PT or unipolar depression, or for screening for modulators, comprises a
 PT BAD-associated protein isoform -

PS Claim 8; Page 34; 163pp; English.

CC The invention relates to a preparation comprising an isolated Bipolar
 CC Affected Disorder (BAD)-Associated Protein Isoform (DPIs). The DPI's are
 CC used to screen, diagnose or prognose of BAD or unipolar depression,
 CC determine the stage or severity of BAD or unipolar depression, identify a
 CC subject at risk of developing BAD or unipolar depression, or monitor the
 CC effect of therapy in a subject. They are also used to screen for or
 CC identify agents that interact with a DPI. These agents, antibodies
 CC against the DPIs, and nucleic acids encoding the DPIs are used to treat
 CC or prevent BAD or unipolar depression. Diseases that can be treated are
 CC attention deficient disorder, a schizoaffective disorder, a bipolar or a
 CC unipolar affective disorder. The DPIs are used in proteomics. The
 CC proteomic approach of using DPIs for screening, diagnosis or prognosis of
 CC BAD or unipolar depression overcomes the problems of using gene
 CC expression analysis, such as not being able to obtain central nervous
 CC system (CNS) tissue from a living patient under normal circumstances.
 CC The present sequence is a DIP decreased in the CSF (cerebro-spinal
 CC fluid) of subjects having BAD.

XX Sequence 7 AA:

Query Match 60.6%; Score 20; DB 22; Length 7;

Best Local Similarity 42.9%; Pred. No. 6.4e+05;

Matches 3; Conservative 3; Mismatches 1; Indels 0; gaps 0;
 OY '1 FTLKISR 7
 * : | : : | |
 Db 1 ytfelst 7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:26:14 ; Search time 40.89 Seconds
(without alignments)
4.181 Million cell updates/sec

Title: US-09-712-819A-6

Perfect score: 33
Sequence: 1 FTLKISR 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 39160

Minimum DB seq length: 0
Maximum DB seq length: 7

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	60.6	7	1	US-08-222-619-8
2	19	57.6	7	5	PCR-US95-04075-8
3	18	54.5	7	6	5252328-8
4	18	54.5	7	1	US-08-346-333-24
5	18	54.5	7	4	US-09-268-992-78
6	18	54.5	7	5	PCT-US91-07506-24
7	17	51.5	7	2	US-09-174-060-17
8	17	51.5	7	3	US-08-654-623-69
9	17	51.5	7	3	US-08-338-382-17
10	16	48.5	5	4	US-08-591-632-23
11	16	48.5	6	1	US-07-717-331F-5
12	16	48.5	7	1	US-07-634-641-12
13	16	48.5	7	1	US-08-136-743B-55
14	16	48.5	7	3	US-09-040-216-28
15	16	48.5	7	4	US-09-173-941-72
16	15	45.5	4	1	US-08-305-871A-17
17	15	45.5	5	1	US-08-136-743B-63
18	15	45.5	5	3	US-09-040-216-55
19	15	45.5	5	4	US-08-591-632-17
20	15	45.5	5	4	US-08-591-632-26
21	15	45.5	6	1	US-08-136-743B-62
22	15	45.5	6	1	US-08-357-264-6
23	15	45.5	6	1	US-08-297-731-4
24	15	45.5	6	1	US-08-222-619-21
25	15	45.5	6	1	US-08-290-448A-41
26	15	45.5	6	1	US-08-672-514-6
27	15	45.5	6	1	US-08-290-448A-41

28	15	45.5	6	1	US-08-175-069A-41	Sequence 41, Appl
29	15	45.5	6	2	US-08-637-759B-219	Sequence 219, App
30	15	45.5	6	3	US-08-871-355A-219	Sequence 219, App
31	15	45.5	6	3	US-09-040-216-54	Sequence 54, Appl
32	15	45.5	6	4	US-09-623-618B-6	Sequence 6, Appl1
33	15	45.5	6	4	US-08-461-939B-41	Sequence 41, Appl
34	15	45.5	6	4	US-08-464-000-41	Sequence 41, Appl
35	15	45.5	6	4	US-09-201-945-219	Sequence 219, App
36	15	45.5	6	4	US-09-187-859-650	Sequence 650, App
37	15	45.5	6	4	US-09-187-859-2845	Sequence 2845, Ap
38	15	45.5	6	4	US-09-187-859-2908	Sequence 2908, Ap
39	15	45.5	6	4	US-09-187-859-3174	Sequence 3174, Ap
40	15	45.5	6	4	US-08-187-859-3264	Sequence 3264, Ap
41	15	45.5	6	5	PCT-US95-04075-21	Sequence 21, Appl
42	15	45.5	6	5	PCT-US95-10793-4	Sequence 4, Appl1
43	15	45.5	6	6	5342922-11	Sequence 6, Appl1
44	15	45.5	7	1	US-08-136-743B-6	Sequence 26, Appl
45	15	45.5	7	1	US-08-136-743B-26	

ALIGNMENTS

RESULT 1
US-08-222-619-8
Sequence 8, Application US/08222619
Patent No. 5652352
GENERAL INFORMATION:
APPLICANT: Lichenstein, Henri
APPLICANT: Lyons, David
APPLICANT: Wurfel, Mark
APPLICANT: Wright, Samuel
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,619
FILING DATE:
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-222-619-8

Query Match 60.6%; Score 20; DB 1; Length 7;
Best local Similarity 57.1%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FTLKISR 7
Db 1 FTFEYSR 7

RESULT 2
PCT-US95-04075-8
Sequence 8, Application PC/TUS9504075

```

; GENERAL INFORMATION:
; APPLICANT: AMGEN INC.
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04075
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; PCT-US95-04075-8

Query Match          60.6%: Score 20; DB 5; Length 7;
Best Local Similarity 57.1%: Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FTLKISR 7
   ||: ||
Db 1 FTFEYSR 7

RESULT 3
5252328-8
; Patent No. 5252328
; APPLICANT: FAULDS, DARYL; VISHOOT, MIMI; BROOKS, EMILY
; TITLE OF INVENTION: MYCOPLASMA HYOPNEUMONIAE ANTIGEN AND USRS
; THEREFORE
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,726
; FILING DATE: 07-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 30,130
; FILING DATE: 26-MAR-1987
; SEQ ID NO: 8:
; LENGTH: 6
; 5252328-8

Query Match          57.6%: Score 19; DB 6; Length 6;
Best Local Similarity 80.0%: Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKI 5
   |||
Db 2 FVLKI 6

RESULT 4
US-08-346-333-24
; Sequence 24 Application US/08346333
; Patent No. 3677153
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
```

```

; APPLICANT: Palzkill, Timothy
; TITLE OF INVENTION: Methods for modifying DNA and for
; TITLE OF INVENTION: detecting effects of such modification on interaction of
; TITLE OF INVENTION: encoded modified polypeptides with target substrates.
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartlin
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,333
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/08/039,501
; FILING DATE:
; APPLICATION NUMBER: US 07/602,158
; FILING DATE: 22-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartlin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-53469/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-346-333-24

Query Match          54.5%: Score 18; DB 1; Length 7;
Best Local Similarity 100.0%: Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTKI 5
   ||||
Db 2 TTKI 5

RESULT 5
US-09-268-992-78
; Sequence 78 Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
```

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 78
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens
US-09-268-992-78

Query Match 54.5%; Score 18; DB 4; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTLKIS 6
:|:|
Db 2 YTMKYS 7

RESULT 6
PCT-US91-07506-24
Sequence 24, Application PC/TUS9107506

GENERAL INFORMATION:
APPLICANT: Botstein, David
APPLICANT: Palzkill, Timothy
TITLE OF INVENTION: Methods for modifying DNA and for
TITLE OF INVENTION: detecting effects of such modification on interaction of
TITLE OF INVENTION: encoded modified polypeptides with target substrates.
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard F. Treccartin
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/07506
FILING DATE: 19911021
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: FP-53469-PC/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US91-07506-24

Query Match 54.5%; Score 18; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TLKI 5
:|:|
Db 2 TLKI 5

RESULT 7
US-09-174-060-17
Sequence 17, Application US/09174060
Patent No. 5989554

GENERAL INFORMATION:
APPLICANT: Knuth, Mark W
APPLICANT: Haak-Frendscho, Mary
APPLICANT: Shultz, John W
APPLICANT: Lesley, Scott A
APPLICANT: Villars, Catherine E
TITLE OF INVENTION: HIGH LEVEL EXPRESSION AND FACILE
TITLE OF INVENTION: PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR
TITLE OF INVENTION: IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ross & Stevens, S.C.
STREET: 1 South Pinckney St.
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/174,060
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/338,382
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 34506.024
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-257-5353
TELEFAX: 608-257-9175
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-174-060-17

Query Match 51.5%; Score 17; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLK 4
:|:|
Db 1 YTLK 4

RESULT 8
US-08-654-623-69
Sequence 69, Application US/08654623
Patent No. 6010884

GENERAL INFORMATION:
APPLICANT: Griffiths, Andrew D
APPLICANT: Holliger, Kaspar-Philipp
APPLICANT: Nissim, Ahuva
APPLICANT: Fisch, Igor
APPLICANT: Winter, Gregory P
TITLE OF INVENTION: Recombinant Binding Proteins and Peptides
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/654,623
; FILING DATE: 29-MAY-1996
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: (C12N 1/21, C12R 1:19)
; FILING DATE: 04-DEC-1992
; APPLICATION NUMBER: GB 9225453.1
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9300816.7
; FILING DATE: 16-JAN-1993
; APPLICATION NUMBER: EP 93303614.7
; FILING DATE: 10-MAY-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9319969.3
; FILING DATE: 22-SEP-1993
; APPLICATION NUMBER: PCT/GB93/02492
; FILING DATE: 03-DEC-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 9412147.2
; FILING DATE: 17-JUN-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/02862
; FILING DATE: 05-DEC-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/448,418
; FILING DATE: 02-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: David W. Clough
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28111/33259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-654-623-69

Query Match          51.5%; Score 17; DB 3; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 TTAKS 6
Db      1 SLKVS 5

RESULT 9
US-08-338-382-17
; Sequence 17, Application US/08338382
; Patent No. 6069230
; GENERAL INFORMATION:
; APPLICANT: Knuth, Mark W
; APPLICANT: Haak-Frendscho, Mary
; APPLICANT: Shultz, John W
; APPLICANT: Lesley, Scott A
; APPLICANT: Villars, Catherine E
; TITLE OF INVENTION: HIGH LEVEL EXPRESSION AND FACILE
; TITLE OF INVENTION: PURIFICATION OF PROTEINS, PEPTIDES AND CONJUGATES FOR
; TITLE OF INVENTION: IMMUNIZATION, PURIFICATION AND DETECTION APPLICATIONS
; NUMBER OF SEQUENCES: 22
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ross & Stevens, S.C.
; STREET: 1 South Pinkney St.
; CITY: Madison
; STATE: WI
; COUNTRY: USA
; ZIP: 53701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/338,382
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sara, Charles S
; REGISTRATION NUMBER: 30,492
; REFERENCE/DOCKET NUMBER: 34506,024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-257-5353
; TELEFAX: 608-257-9175
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-338-382-17

Query Match          51.5%; Score 17; DB 3; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 FTLK 4
Db      1 YTLK 4

RESULT 10
US-08-591-632-23
; Sequence 23, Application US/08591632
; Patent No. 6261558
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; APPLICANT: Burton, Dennis R.
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
; TITLE OF INVENTION: ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10550 No. 6261558th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,632
; FILING DATE: 19-OCT-1994
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/11907
; FILING DATE: 19-OCT-1994
; PRIORITY APPLICATION DATA:
```

APPLICATION NUMBER: US 08/308,841
FILING DATE: 19-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/233,619
FILING DATE: 26-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/139,409
FILING DATE: 19-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 332.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 784-2937
TELEFAX: (619) 784-9399
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-591-632-23

Query Match 48.5%; Score 16; DB 4; Length 5;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLLK 4
DB 2 FTLLQ 5

RESULT 11
US-07-717-331F-5
Sequence 5, Application US/07717331F
Patent No. 5484905
GENERAL INFORMATION:
APPLICANT: June Nasrallah; Michael Nasrallah; and Joshua
APPLICANT: Stein
TITLE OF INVENTION: A Receptor Protein Kinase Gene
TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: USA
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/717,331F
FILING DATE: June 19th 1991
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-717-331F-5

Query Match 48.5%; Score 16; DB 1; Length 6;
Best Local Similarity 75.0%; Pred. No. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 LKIS 6
DB 2 LKVS 5

RESULT 12
US-07-634-641-12
Sequence 12, Application US/07634641
Patent No. 5386011
GENERAL INFORMATION:
APPLICANT: Wiedeman, Paul E.
APPLICANT: Kawal, Megumi
APPLICANT: Luly, Jay R.
APPLICANT: Or, Yal-Sun
APPLICANT: Wagner, Rolf
TITLE OF INVENTION: Hexa- and Heptapeptide Anaphylatoxin
TITLE OF INVENTION: Receptor Ligands
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: One Abbott Park Road
CITY: No. 5386011th Chicago
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/634,641
FILING DATE: 19901227
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Janssen, Jerry F.
REGISTRATION NUMBER: 29,175
REFERENCE/DOCKET NUMBER: 4934.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (708) 937-4558
TELEFAX: (708) 937-9556
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-634-641-12

Query Match 48.5%; Score 16; DB 1; Length 7;
Best Local Similarity 28.6%; Pred. No. 1.7e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FTLLKSR 7
DB 1 FRMRLLGR 7

RESULT 13
US-08-136-743B-55
Sequence 55, Application US/08136743B
Patent No. 5459063
GENERAL INFORMATION:
APPLICANT: Barry S. Cooperman, Harvey Rubin,
APPLICANT: Jerome Salem, and Allison L. Fisher
TITLE OF INVENTION: "Plasmodium falciparum Ribonu-

```

; TITLE OF INVENTION: clectide Reductase, DNA Sequences Therefor and Peptide Inhibi
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The University of Pennsylvania
; STREET: Suite 330
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19104-3246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,743B
; FILING DATE: 10/14/93
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 3957-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549
; TELEX: NO. 5459063e
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-136-743B-55

Query Match          48.5%; Score 16; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. NO. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKI 5
   |||
Db 1 FTLDL 5

RESULT 14
US-09-040-216-28
; Sequence 28, Application US/09040216
; Patent No. 6030942
; GENERAL INFORMATION:
; APPLICANT: COOPERMAN, ET AL., BARRY
; TITLE OF INVENTION: PEPTIDES, PEPTIDE ANALOGS, PEPTIDOMIMETICS, AND OTHER
; TITLE OF INVENTION: SMALL MOLECULES USEFUL FOR INHIBITING THE ACTIVITY OF
; TITLE OF INVENTION: RIBONUCLEOTIDE REDUCTASE
; FILE REFERENCE: 9596-63U1
; CURRENT APPLICATION NUMBER: US/09/040,216
; CURRENT FILING DATE: 1998-03-17
; EARLIER APPLICATION NUMBER: 08/919,748
; EARLIER FILING DATE: 1997-08-28
; EARLIER APPLICATION NUMBER: 60/025,146
; EARLIER FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Ribonucleotide reductase inhibitor peptide
; FEATURE:
; OTHER INFORMATION: residue 1: MOD_RES: ACETYLATION
;
US-09-040-216-28
```

```

Query Match          48.5%; Score 16; DB 3; Length 7;
Best Local Similarity 60.0%; Pred. NO. 1.7e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKI 5
   |||
Db 1 FTLDL 5
```

```

RESULT 15
US-09-173-941-72
; Sequence 72, Application US/09173941
; Patent No. 6140081
; GENERAL INFORMATION:
; APPLICANT: BARBAS, Carlos F.
; TITLE OF INVENTION: ZINC FINGER BINDING DOMAINS FOR GNN
; FILE REFERENCE: NOV00815
; CURRENT APPLICATION NUMBER: US/09/173,941
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: nucleotide
;
US-09-173-941-72
```

```

Query Match          48.5%; Score 16; DB 4; Length 7;
Best Local Similarity 75.0%; Pred. NO. 1.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 KISR 7
   |||
Db 4 KLSR 7
```

Search completed: July 15, 2002, 13:26:15
job time: 495 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:41:35 : Search time 14.1 Seconds

(without alignments)
61.334 Million cell updates/sec

Title: US-09-712-819A-7

Perfect score: 43

Sequence: 1 TDFLTRSS 9

Scoring table:

BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 28338 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 788

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	51.2	8	2	T14906 hypothetical prote
2	16	37.2	9	2	A23967 leucopyrokinin - M
3	16	37.2	9	2	A28924 fructose-bisphosph
4	14	32.6	8	2	S71919 alcohol dehydrogen
5	14	32.6	9	2	A24244 adipokineic hormo
6	14	32.6	9	2	G85802 hypothetical prote
7	13	30.2	6	2	PD0028 pcy-kinin 2 - pena
8	13	30.2	8	2	S11545 adipokineic hormo
9	13	30.2	8	2	S13661 polygalacturonase
10	13	30.2	9	2	B38740 Ig kappa chain C r
11	12	27.9	5	2	A32516 T-cell receptor be
12	12	27.9	5	2	PT0644 cholecystokinin-5
13	12	27.9	7	2	B39127 phosphotransferase
14	12	27.9	7	2	E30608 Ig kappa chain V-I
15	12	27.9	7	2	PT0665 T-cell receptor be
16	12	27.9	8	2	S10596 adipokineic hormo
17	12	27.9	8	2	S55310 adipokineic hormo
18	12	27.9	8	2	A58620 cholecystokinin -
19	12	27.9	8	2	PD0012 cholecystokinin -
20	12	27.9	8	2	A43001 cholecystokinin -
21	12	27.9	8	2	PL0184 capsid protein VP-
22	12	27.9	8	2	US0318 leucokinin VIII -
23	12	27.9	8	2	H41978 callifMRamide 8 -
24	12	27.9	8	2	E47393 neuropeptide calla
25	12	27.9	8	2	S65381 cytochrome-c oxida
26	12	27.9	8	2	S66296 Na+-transporting A
27	12	27.9	9	2	A61357 phyllocaerulein -
28	12	27.9	9	2	A43848 cell surface adhes
29	12	27.9	9	2	S36898 ribosomal protein

30	12	27.9	9	2	A41978 callifMRamide 1 -
31	12	27.9	9	2	A44787 callifMRamide 10
32	12	27.9	9	2	B41978 callifMRamide 2 -
33	12	27.9	9	2	callifMRamide 3 -
34	12	27.9	9	2	callifMRamide 4 -
35	12	27.9	9	2	callifMRamide 5 -
36	12	27.9	9	2	F41978 callifMRamide 6 -
37	12	27.9	9	2	callifMRamide 7 -
38	12	27.9	9	2	A45199 L-hyosophorin - Ja
39	12	27.9	9	2	PT0299 Ig heavy chain CRD
40	11	25.6	4	2	I40697 biotin A - Citroba
41	11	25.6	4	2	A40135 branched-chain-aml
42	11	25.6	5	2	B44817 34.5k structural p
43	11	25.6	5	2	D44817 35k structural pro
44	11	25.6	8	2	A32523 peptidyl-dipeptida
45	11	25.6	8	2	S08996 hypertrehalosemic

ALIGNMENTS

RESULT 1
T14906
hypothetical protein - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14906
R:Feidbrugge, M.; Sprenger, M.; Dinkelbach, M.; Yazaki, K.; Harter, K.; Weisshaar, B.
Plant Cell 6, 1607-1621, 1994
A:Title: Functional analysis of a light-responsive plant bZIP transcriptional regulat
A:Reference number: 218259; MUID:95128172
A:Accession: T14906
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-8 <FEL>
A:Cross-references: EMBL:S75395; NID:9913201; PID:6194245

Query Match 51.2%; Score 22; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TLTIS 8
Db 4 TLTIS 8

RESULT 2
A23967
leucopyrokinin - Madeira cockroach
C:Species: Leucophaea maderae (Madeira cockroach)
C:Date: 31-Mar-1988 #sequence_revision 26-May-1994 #text_change 11-Jul-1997
C:Accession: A23967
R:Nachman, R.J.; Holman, G.M.; Cook, B.J.
Biochem. Biophys. Res. Commun. 137, 936-942, 1986
A:Title: Active fragments and analogs of the insect neuropeptide leucopyrokinin: stru
A:Reference number: A23967; MUID:66265041
A:Accession: A23967
A:Molecule type: protein
A:Residues: 1-8 <MAC>
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F1/Modified site: pyrrolidone carboxyl acid (Gln) #status experimental
F8/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 37.2%; Score 16; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TDFPT 4
Db 2 TSFT 5

```
RESULT 3
A26924
fructose-bisphosphate aldolase (EC 4.1.2.13) B. hepatic - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 07-Feb-1997
C:Accession: A26924
R:Lacko, A.G.; Brox, L.W.; Gracy, R.W.; Horecker, B.L.
J. Biol. Chem. 245, 2140-2141, 1970
A>Title: The carboxyl-terminal structure of rabbit liver aldolase (aldolase B).
A:Reference number: A26924; MUID:70166720
A:Accession: A26924
A:Molecule type: protein
A:Residues: 1-9 <LAK>
C:Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis; liver; per

Query Match      37.2%; Score 16; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 FTLTRSS 9
   | 1 1 1
Db 2 FLATTSS 8

RESULT 4
S71919
alcohol dehydrogenase (EC 1.1.1.1) - grass carp (fragment)
C:Species: Ctenopharyngodon idella (grass carp)
C>Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 07-May-1999
C:Accession: S71919
R:Tsuji, H.T.; Mock, W.Y.; Lau, K.K.; Fong, W.P.
Biochim. Biophys. Acta 1296, 41-46, 1996
A>Title: Proteolytic activation of grass carp (Ctenopharyngodon idellus) liver alcohol de
A:Reference number: S71919; MUID:96350418
A:Accession: S71919
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <TSU>
A>Note: the source is designated Ctenopharyngodon idellus
C:Keywords: NAD; oxidoreductase

Query Match      32.6%; Score 14; DB 2; Length 8;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDFT 4
   | 1 1
Db 5 TRFT 8

RESULT 5
A24244
adipokinetic hormone - bollworm
N:Alternate names: Hez-AKH
C:Species: Heliothis zea (bollworm, corn earworm, tomato fruitworm)
C>Date: 31-Mar-1988 #sequence_revision 23-Mar-1995 #text_change 31-Oct-1997
C:Accession: A24244
R:Jafté, H.; Ralua, A.K.; Riley, C.T.; Fraser, B.A.; Holman, G.M.; Wagner, R.M.; Ridgway
Blochem. Biophys. Res. Commun. 135, 622-628, 1986
A>Title: Isolation and primary structure of a peptide from the corpora cardiaca of Helio
A:Reference number: A24244; MUID:86186794
A:Accession: A24244
A:Molecule type: protein
A:Residues: 1-9 <JAF>
C:Superfamily: adipokinetic hormone
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:9/Modified site: amidated carboxyl end (Gly) #status experimental
```

```
Query Match      32.6%; Score 14; DB 2; Length 9;
Best Local Similarity 60.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 LTRSS 9
   | 1 1
Db 2 LTRFS 6

RESULT 6
G85802
hypothetical protein 22947 [imported] - Escherichia coli (strain O157:H7, substrain E
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85802
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lm, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206531
A:Accession: G85802
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-9 <STO>
A:Cross-references: GB:AE005174; NID:q12515957; PIDN:AXG56883.1; GSPDB:GN00145; OMGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 22947
```

```
Query Match      32.6%; Score 14; DB 2; Length 9;
Best Local Similarity 28.6%; Pred. No. 2.8e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 FTLTRSS 9
   | 1 1 1
Db 3 YTFMLSN 9

RESULT 7
PD0028
pey-kinin 2 - penaeid shrimp (Penaeus vannamei) (fragment)
C:Species: Penaeus vannamei
C>Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 19-May-2000
C:Accession: PD0028
R:Nieco, J.; Veelaert, D.; Derna, R.; Waelkens, E.; Cerslaens, A.; Coast, G.; Devree
Blochem. Biophys. Res. Commun. 248, 406-411, 1998
A>Title: Identification of one tachykinin- and two kinin-related peptides in the brai
A:Reference number: PD0027; MUID:98342103
A:Accession: PD0028
A:Molecule type: protein
A:Residues: 1-6 <NIE>
C:Comment: This peptide belongs to myotropic neuropeptides.
```

```
Query Match      30.2%; Score 13; DB 2; Length 6;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPT 4
   | 1 1
Db 1 DPT 3

RESULT 8
S11545
adipokinetic hormone - nestling-sucking blowfly
C:Species: Protophormia terrenovae (nestling-sucking blowfly)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 31-Oct-1997
C:Accession: S11545
R:Gaede, G.; Milps, H.; Kellner, R.
Blochem. J. 269, 309-313, 1990
A>Title: Isolation and structure of a novel charged member of the red-pigment-concent
```

erraeenqvae (Diptera).

A:Reference number: S11545; MUID:90351345

A:Accession: S11545

A:Molecule type: protein

A:Residues: 1-8 <GAE>

C:Superfamily: adpokinetic hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamid

F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental

F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match

Best Local Similarity 30.2%; Score 13; DB 2; Length 8;

Best Local Similarity 75.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 LRTS 8

DB 2 LRTS 5

RESULT 9

S13661 polygalacturonase (EC 3.2.1.15) isoform PG2 - fungus (Sclerotinia sclerotiorum) (fragment)

N:Alternate names: endopolygalacturonase; pectin depolymerase; pectinase

C:Species: Sclerotinia sclerotiorum

C:Date: 19-Mar-1997 #sequence_revision 21-Nov-1998 #text_change 21-Nov-1998

C:Accession: S13661

R:Wakeman, G.; Keon, J.P.R.; Turner, G.

Biochim. Biophys. Acta 1073, 43-48, 1991

A:Title: Purification and characterization of two endopolygalacturonases from Sclerotinia

A:Reference number: S13661; MUID:91120822

A:Accession: S13661

A:Molecule type: protein

A:Residues: 1-8 <NAK>

C:Function:

A:Description: involved in pectin degradation

C:Keywords: glycosidase; hydrolase

Query Match

Best Local Similarity 30.2%; Score 13; DB 2; Length 8;

Best Local Similarity 60.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TLTIS 8

DB 3 TLTIS 7

RESULT 10

B38740 I9 kappa chain C region (PY20) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998

C:Accession: B38740

R:Ruff-Jamison, S.; Campos-Gonzalez, R.; Glenney Jr., J.R.

J. Biol. Chem. 266, 6607-6613, 1991

A:Title: Heavy and light chain variable region sequences and antibody properties of anti

A:Reference number: A38740; MUID:91177923

A:Accession: B38740

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-9 <RUP>

Query Match

Best Local Similarity 30.2%; Score 13; DB 2; Length 9;

Best Local Similarity 37.5%; Pred. No. 2.8e+05;

Matches 3; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 DFTLTIS 9

DB 2 DAAPTVGS 9

RESULT 11

A32516 cholecystokinin-5 - dog

N:Alternate names: CCK-5

C:Species: Canis lupus familiaris (dog)

C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 18-Aug-2000

C:Accession: A32516

R:Shively, J.; Reeve Jr., J.R.; Eysselein, V.E.; Ben-Avram, C.; Vigna, S.R.; Walsh, J

Am. J. Physiol. 252, G272-G275, 1987

A:Title: CCK-5: sequence analysis of a small cholecystokinin from canine brain and in

A:Reference number: A32516; MUID:87153871

A:Accession: A32516

A:Molecule type: protein

A:Residues: 1-5 <SHD>

C:Comment: This peptide corresponds to the five carboxyl-terminal residues of cholecy

C:Superfamily: gastrin

C:Keywords: amidated carboxyl end; neuropeptide

F:5/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match

Best Local Similarity 27.9%; Score 12; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DF 3

DB 4 DF 5

RESULT 12

PT0644 T-cell receptor beta chain V-D-J region (111-16) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997

C:Accession: PT0644

R:Feeney, A.J.

J. Exp. Med. 174, 115-124, 1991

A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions

A:Reference number: PT0509; MUID:91277601

A:Accession: PT0644

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-5 <FEE>

A:Experimental source: newborn thymus, strain BALB/c

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 27.9%; Score 12; DB 2; Length 5;

Best Local Similarity 50.0%; Pred. No. 2.8e+05;

Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDF 4

DB 2 SSFT 5

RESULT 13

B39127 phosphotransferase system enzyme II (EC 2.7.1.65) - Escherichia coli (fragment)

C:Species: Escherichia coli

C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999

C:Accession: B39127

R:Hardesty, C.; Ferran, C.; Di Rienzo, J.M.

J. Bacteriol. 173, 449-456, 1991

A:Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of

A:Reference number: A39127; MUID:91100329

A:Accession: B39127

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-7 <HAR>

A:Cross-references: GB:M38416; NID:9155142; PID:AA98418.1; PID:9155144

C:Keywords: phosphotransferase

Query Match 27.9%; Score 12; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0;

QY 2 DF 3
 ||
 DB 2 DF 3

RESULT 14

E30608
 I9 Kappa chain V-III region (Cag) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 29-Jun-1989 #sequence_revision 29-Jun-1989 #text_change 16-Aug-1996
 C:Accession: E30608
 R:Contl. F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Sold
 J. Immunol. 142, 3158-3163, 1989
 A:Title: Structural and idiotype characterization of the L chains of human IgM autoantib
 A:Reference number: A30601; MUID:89215279
 A:Accession: E30608
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-7 <CON>
 C:Keywords: heterotrimer; immunoglobulin

Query Match 27.9%; Score 12; DB 2; Length 7;
 Best Local Similarity 42.9%; Pred. No. 2.8e+05;
 Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 DFTLITS 8
 : || |
 DB 1 EIVLTQS 7

RESULT 15

PT0665
 T-cell receptor beta chain V-D-J region (121-38M) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
 C:Accession: PT0665
 R:Feeney, A.J.
 J. Exp. Med. 174, 115-124, 1991
 A:Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
 A:Reference number: PT0509; MUID:91277601
 A:Accession: PT0665
 A>Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-7 <FEES>
 A:Experimental source: day 4 postnatal thymus, strain BALB/c
 C:Keywords: T-cell receptor

Query Match 27.9%; Score 12; DB 2; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DF 3
 ||
 DB 6 DF 7

Search completed: July 15, 2002, 13:43:36
 Job time: 121 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:43:21; Search time 10.19 Seconds

(Without alignments)
34.198 Million cell updates/sec

Title: US-09-712-819a-7

Perfect score: 43

Sequence: 1 TDPFTLRIS 9

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 231

Minimum DB seq length: 0

Maximum DB seq length: 9

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	17	39.5	8 1	LMT2_LOCM1
2	17	39.5	9 1	UPA3_HUMAN
3	16	37.2	8 1	LPR_LEUMA
4	15	34.9	9 1	UHA2_HUMAN
5	14	32.6	7 1	AL17_CYDPO
6	14	32.6	8 1	AL15_CYDPO
7	14	32.6	8 1	CAD1_ENTFA
8	14	32.6	9 1	DL_NBPNO
9	12	27.9	5 1	RE11_LITRU
10	12	27.9	7 1	GRRP_MOUSE
11	12	27.9	8 1	AKH_LTBAB
12	12	27.9	8 1	AL15_CALVO
13	12	27.9	8 1	CCRN_MACEU
14	12	27.9	8 1	COXG_RAT
15	12	27.9	8 1	FAR8_CALVO
16	12	27.9	8 1	FUSS_FUSSO
17	12	27.9	8 1	LCK8_LEUMA
18	12	27.9	9 1	FAR1_CALVO
19	12	27.9	9 1	FAR2_CALVO
20	12	27.9	9 1	FAR3_CALVO
21	12	27.9	9 1	FAR4_CALVO
22	12	27.9	9 1	FAR5_CALVO
23	12	27.9	9 1	FAR6_CALVO
24	12	27.9	9 1	FAR7_CALVO
25	12	27.9	9 1	FAR8_CALVO
26	11	25.6	5 1	BIOA_CITFR
27	11	25.6	5 1	RE31_LITRU
28	11	25.6	5 1	RE32_LITRU
29	11	25.6	8 1	AKH_TABAT
30	11	25.6	8 1	HTF2_PERAM
31	11	25.6	8 1	ORMY_ORCLI
32	11	25.6	9 1	CCAP_CARMA
33	11	25.6	9 1	PGIR_DIAAB

34	10	23.3	5 1	ET04_LITRU
35	10	23.3	8 1	LPMS_STREP
36	10	23.3	9 1	LPCH_STAAU
37	9	20.9	7 1	CCF1_ENTFA
38	9	20.9	7 1	CIA_ENTFA
39	9	20.9	8 1	AL11_CYDPO
40	9	20.9	8 1	AL16_CYDPO
41	9	20.9	8 1	UPA3_HUMAN
42	9	20.9	9 1	BUR_CLOPA
43	9	20.9	9 1	ISOF_CYCA
44	9	20.9	9 1	MGMT_BOVIN
45	9	20.9	1	NSK1_SARBU

ALIGNMENTS

RESULT 1				
LMT2_LOCM1	STANDARD:	PRT:	8 AA.	
ID LMT2_LOCM1				
AC P22396;				
DT 01-AUG-1991 (Rel. 19, Created)				
DT 01-AUG-1991 (Rel. 19, Last sequence update)				
DT 01-AUG-1991 (Rel. 19, Last annotation update)				
DE Locustamyotropin 2 (LOM-WT-2).				
OS Locusta migratoria (Migratory locust).				
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;				
OC Acridoidea; Acridoidea; Acrididae; Locusta.				
OX NCBI_TaxID=7004;				
RN [1]				
RP SEQUENCE.				
RC TISSUE=Corpora cardiaca;				
RA Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;				
RT "Isolation, identification and synthesis of locustamyotropin II, an				
RT additional neuropeptide of locusta migratoria. Member of the				
RT cephalomyotrophic peptide family.";				
RL Insect Biochem. 20:479-484(1990).				
CC -I- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY				
CC -I- (MYOTROPIC ACTIVITY).				
CC -I- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.				
DR InterPro: IPR001484; PYROKININ.				
DR PROSITE: PS00539; PYROKININ: 1.				
KW Neuropeptide; Amidation; Pyrokinin.				
FT MODRES 8 AA: 934 MW: 26341771A9CA87B CRC64;				
SO SEQUENCE				
Query Match 39.5%; Score 17; DB 1; Length 8;				
Best Local Similarity 100.0%; Pred. No. 1e+05;				
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY 2 DFT 4				
DB 3 DFT 5				
RESULT 2				
UPA3_HUMAN	STANDARD:	PRT:	9 AA.	
ID UPA3_HUMAN				
AC P30089;				
DT 01-APR-1993 (Rel. 25, Created)				
DT 01-APR-1993 (Rel. 25, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DE Unknown protein from 2D-page of plasma (Spot 11) (Fragment).				
OS Homo sapiens (Human).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX NCBI_TaxID=9606;				
RN [1]				
RP SEQUENCE.				
RC TISSUE=Plasma;				
RX MEDLINE=93092937; PubMed=1459097;				

RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquall C.,
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
 RA Hochstrasser D.F.;
 RT "Plasma protein map: an update by microsequencing.";
 RL Electrophoresis 13:707-714(1992).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.6, ITS MW IS: 46 kDa.
 DR SWISS-2DPAGE: P30089; HUMAN.
 FT NON_TER 1 1
 FT MOD_RES 9 9
 SQ SEQUENCE 9 AA: 1056 MW: 26F2B1BAF769C737 CRC64;

Query Match 39.5%; Score 17; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDF 3
 DB 7 TDF 9

RESULT 3
 LPK_LEUMA STANDARD: PRT; 8 AA.
 ID LPK_LEUMA
 AC P13049;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Leukopyrokinin (LPK) (LEW-PK).
 OS Leucophaea maderae (Madelera cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberoidea; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86269041; PubMed=3015140;
 RA Nachman R.J., Holman G.M., Cook B.J.;
 RT "Active fragments and analogs of the insect neuropeptide
 RL leucopyrokinin: structure-function studies.";
 RL Biochem. Biophys. Res. Commun. 137:936-942(1986).
 RN [2]
 RP SEQUENCE AND SYNTHESIS.
 RC TISSUE-Head;
 RX MEDLINE=87052651; PubMed=2877794;
 RA Holman G.M., Cook B.J., Nachman R.J.;
 RT "Primary structure and synthesis of a blocked myotropic
 RT neuropeptide isolated from the cockroach, Leucophaea maderae.";
 RL Comp. Biochem. Physiol. 85C:219-224(1986).
 CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
 CC (MYOTROPIC ACTIVITY).
 CC -1- MISCELLANEOUS: AN ANALOG WITHOUT THE N-TERMINAL PCA RESIDUE WAS
 CC SYNTHESIZED AND FOUND TO EXHIBIT GREATER ACTIVITY (144%) THAN THE
 CC PARENT NEUROPEPTIDE. THE PORTION OF THE SEQUENCE OF LPK MOST
 CC CRITICAL FOR THE MYOTROPIC PROPERTIES IS LIMITED TO THE
 CC PENTAPEPTIDE FRAGMENT FTPL.
 CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
 DR InterPro: IPR001484; Pyrokinin.
 DR PROSITE: PS00539; PYROKININ: 1.
 KW Neuropeptide; Amidation; Pyrokinin.
 FT MOD_RES 1 1
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA: 949 MW: 92341771A9D5A1B6 CRC64;

Query Match 37.2%; Score 16; DB 1; Length 8;
 Best Local Similarity 75.0%; Pred. NO. 1e+05;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDF 4
 DB 1 TDF 1

DB 2 TSFT 5

RESULT 4
 ID UHA2_HUMAN STANDARD: PRT; 9 AA.
 AC P40929;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of heart (Spot 5603) (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Heart;
 RX MEDLINE=95203287; PubMed=7895732;
 RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
 RT "The human myocardial two-dimensional gel protein database: update
 RT 1994.";
 RL Electrophoresis 15:1459-1465(1994).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.0, ITS MW IS: 55.3 kDa.
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA: 1104 MW: 8874B1B5B01B2CA CRC64;

Query Match 34.9%; Score 15; DB 1; Length 9;
 Best Local Similarity 50.0%; Pred. NO. 1e+05;
 Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDF 4
 DB 4 TDF 7

RESULT 5
 ID ALL7_CYDPO STANDARD: PRT; 7 AA.
 AC P82158;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Cydlaestatin 7.
 OS Cydia pomonella (Coddling moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Tortricoidae; Tortricidae; Olethreutinae; Cydia.
 OX NCBI_TaxID=82600;
 RN [1]
 RP SEQUENCE.
 RC TISSUE-Larva;
 RX MEDLINE=98054539; PubMed=9392829;
 RA Duve H., Johnsen A.H., Maestri J.-L., Scott A.G., Winstanley D.,
 RA Davey M., East P.D., Thorpe A.;
 RT "Lepidopteran peptides of the allatostatin superfamily.";
 RL Peptides 18:1301-1309(1997).
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 KW Neuropeptide; Amidation.
 FT MOD_RES 7 7
 SQ SEQUENCE 7 AA: 873 MW: 672879CABB569350 CRC64;

Query Match 32.6%; Score 14; DB 1; Length 7;
 Best Local Similarity 75.0%; Pred. NO. 1e+05;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 DFTL 5
 DB 4 DFTL 7

```

RESULT 6
AL15_CYPDO STANDARD; PRT; 8 AA.
ID
AC P82156;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cydiastatin 5.
OS Cydia pomonella (Coddling moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrystia;
OC Tortricidae; Tortricidae; Olethreutinae; Cydia.
OX NCBI_TaxID=82600;
RN [1]
RP SEQUENCE.
RC TISSUE=Larva;
RX MEDLINE=98054539; PubMed=9392829;
RA Dave H., Johnson A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA Davey M., East P.D., Thorpe A.;
RT "Lepidopteran peptides of the allatostatin superfamily.";
RL Peptides 18:1301-1309(1997).
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KM Neuropeptide; Amidation.
FT MOD_RES 8 AA; 898 MW; 922879CAB858640D CRC64;
SQ SEQUENCE

```

```

Query Match 32.6%; Score 14; DB 1; Length 8;
Best Local Similarity 75.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 DFTL 5
DB 5 DFGL 8

RESULT 7
CADI_ENTFA STANDARD; PRT; 8 AA.
ID
AC P13268;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CADI.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE.
RX MEDLINE=85051889; PubMed=6437872;
RA Morf M., Sasekani Y., Narita M., Isogai A., Fujino M., Kitada C.,
RA Craig R.A., Clewell D.B., Suzuki A.;
RT "Isolation and structure of the bacterial sex pheromone, CADI, that
RT induces plasmid transfer in Streptococcus faecalis.";
RL FEBS Lett. 178:97-100(1984).
CC -1- FUNCTION: CADI IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC HEMOLYSIN PLASMID PADI.
KM Pheromone.
SQ SEQUENCE 8 AA; 819 MW; 047DD732C735B9C7 CRC64;

```

```

Query Match 32.6%; Score 14; DB 1; Length 8;
Best Local Similarity 33.3%; Pred. No. 1e+05;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 3 FTLTIS 8
DB 2 FSLVLA 7

```

```

RESULT 8

```

```

D1_NEPNO STANDARD; PRT; 9 AA.
ID
AC P24816;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE Gastrin/cholecystokinin-Like peptide D1.
OS Nephrops norvegicus (Norway lobster).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC Nephropidae; Nephropidae; Nephrops.
OX NCBI_TaxID=6829;
RN [1]
RP SEQUENCE.
RC TISSUE=Stomach;
RX MEDLINE=92082847; PubMed=1747388;
RA Favrel P., Kegel G., Sedlmeier D., Keller R., van Wormhoudt A.;
RT "Structure and biological activity of crustacean gastrointestinal
RT peptides identified with antibodies to gastrin/cholecystokinin.";
RL Biochimie 73:1233-1239(1991).
CC -1- FUNCTION: MAY CONTROL DIGESTION PROCESSES IN CRUSTACEANS.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR; A48398; A48398.
KM Hormone.
SQ SEQUENCE 9 AA; 1038 MW; 60EC79CAB6D8787B CRC64;

```

```

Query Match 32.6%; Score 14; DB 1; Length 9;
Best Local Similarity 75.0%; Pred. No. 1e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 DFTL 5
DB 6 DFWL 9

```

```

RESULT 9
REIL_LITRU STANDARD; PRT; 5 AA.
ID
AC P82070;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Rubellidin 1.1.
OS Litorea rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litorea.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE. AND MASS SPECTROMETRY.
RC TISSUE=Skin secretion;
RA Steinboerner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA Tyler M.J., Wallace J.C.;
RT "The structure of new peptides from the Australian red tree frog
RT 'Litorea rubella', the skin peptide profile as a probe for the study
RT of evolutionary trends of amphibians.";
RL Aust. J. Chem. 49:955-963(1996).
CC -1- FUNCTION: SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR ANTIBIOTIC
CC ACTIVITY.
CC -1- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC -1- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KM Amphibian skin.
SQ SEQUENCE 5 AA; 598 MW; 6DD9C9CAB2A00000 CRC64;

```

```

Query Match 27.9%; Score 12; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 DF 3
DB 2 DF 3

```

```

RESULT 10
GFRP_MOUSE
ID GFRP_MOUSE STANDARD: PRT; 7 AA.
AC P99025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP cyclohydrolase I feedback regulatory protein (P35) (Fragment).
GN GCHFR OR GFRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RA Sanchez J.-C., Rouge V., Frutiger S., Hughes G., Van J.X.,
RA Hoogland C., Appel R.D., Blinz P.A., Hochstrasser D.F.,
RA Cowhorne M.;
RL Submitted (Aug-1998) to the SWISS-PROT data bank.
CC -1- FUNCTION: MEDIATES TETRAHYDROPIPERIN INHIBITION OF GTP
CC CYCLOHYDROLASE I. THIS INHIBITION IS REVERSED BY L-PHENYLALANINE
CC (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
DR SWISS-2DPAGE; P99025; MOUSE.
FT INIT_MET 0
FT NON_TER 7
SQ SEQUENCE 7 AA; 806 MW; 715B057273B4700 CRC64;

```

```

Query Match 27.9%; Score 12; DB 1; Length 7;
Best Local Similarity 60.0%; Pred. No. 1e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 5 LTRSS 9
DB 3 LIST 7

```

```

RESULT 11
AKH_LIBAU
ID AKH_LIBAU STANDARD: PRT; 8 AA.
AC P23418;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Adipokine hormone (AKH).
OS Libellula auripennis (Skimmer dragonfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Palaeoptera; Odonata; Anisoptera; Libellulidae; Libellula.
NCBI_TaxID=6966;
RN [1]
RP SEQUENCE AND SYNTHESIS.
RC TISSUE=Corpora cardiaca;
RA MEDLINE=90359035; PubMed=2390213;
RA Gaede G.;
RT "The putative ancestral peptide of the adipokine/red-pigment-
RT concentrating hormone family isolated and sequenced from a
RT dragonfly."
RL Biol. Chem. Hoppe-Seyler 371:475-483(1990).
CC -1- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -1- SIMILARITY: BELONGS TO THE AKH / HRH / RCH FAMILY.
DR PIR: S10566; S10596.
DR InterPro: IPR002047; AKH.
DR PROSITE: PS00256; AKH; 1.
RW Neuropeptide; Amidation; flight.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.

```

```

SQ SEQUENCE 8 AA; 978 MW; 8665A771A9C452D6 CRC64;

```

```

Query Match 27.9%; Score 12; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 DFT 4
DB 3 NTF 5

```

```

RESULT 12
ALIS_CALVO
ID ALIS_CALVO STANDARD: PRT; 8 AA.
AC P41841;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Callistostatin 5 (Met-callistostatin 1) ([Hyp3]Met-callistostatin).
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestroidea; Calliphoridae; Calliphora.
NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RX MEDLINE=93211980; PubMed=8460157;
RA Duye H., Johnsen A.H., Scott A.G., Yu C.G., Yagci K.J., Tobe S.S.,
RA Thorpe A.;
RT "Callistostatin: neuropeptides from the blowfly Calliphora vomitoria
RT with sequence homology to cockroach allatostatin."
RL Proc. Natl. Acad. Sci. U.S.A. 90:2456-2460(1993).
RN [2]
RP CHARACTERIZATION AND HYDROXYLATION.
RC TISSUE=Head;
RX MEDLINE=94342269; PubMed=8063725;
RA Duye H., Johnsen A.H., Scott A.G., East P., Thorpe A.;
RT "[Hyp3]Met-callistostatin. Identification and biological properties of
RT a novel neuropeptide from the blowfly Calliphora vomitoria."
RL J. Biol. Chem. 269:21059-21066(1994).
CC -1- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR AND PLAY
CC A ROLE IN THE INTERACTION OF INFORMATION WITHIN THE BRAIN. MAY BE
CC INVOLVED IN THE CONTROL OF VISCERAL MUSCLES DUE TO ITS ABILITY TO
CC BEHAVE AS POTENT INHIBITORS OF PERISTALTIC MOVEMENTS. MAY ALSO
CC FULFILL A NEUROHORMONAL ROLE ON MUSCLES OF THE GUT AND HEART.
CC -1- TISSUE SPECIFICITY: NEURONS WITHIN BRAIN AND ABDOMINAL GANGLION.
CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
DR PIR: E47393; E47393.
DR Neuropeptide; Amidation; Hydroxylation.
RW MOD_RES 3 3 HYDROXYLATION (IN 20% OF THE PEPTIDES).
FT MOD_RES 8 8 AMIDATION.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 883 MW; 7D9879CABBA77768 CRC64;

```

```

Query Match 27.9%; Score 12; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DFT 3
DB 5 DFT 6

```

```

RESULT 13
CCKN_MACEU
ID CCKN_MACEU STANDARD: PRT; 8 AA.
AC P30369;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Cholecystokinin (CCK).
GN CCK.

```


OS Macropus eugenii (Tamar wallaby), and
 OS Dasypus viverrinus (Southeastern quoll).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
 OX NCBI_TaxID=9315, 9279;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=88234141; PubMed=3375140;
 RA Fan Z.W., Eng J., Shaw G., Yalow R.S.;
 RT "Cholecystokinin octapeptide purified from brains of Australian
 marsupials.";
 RL Peptides 9:429-431(1988).
 CC -1- FUNCTION: THIS PEPTIDE HORMONE INDUCES GALL BLADDER CONTRACTION
 AND THE RELEASE OF PANCREATIC ENZYMES IN THE GUT. ITS FUNCTION
 IN THE BRAIN IS NOT CLEAR.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 DR PIR: A43001; A43001.
 DR InterPro: IPR001651; Gastrin.
 DR PROSITE: PS00259; GASTRIN; 1.
 KW Amidation; Sulfation; Hormone.
 FT MOD_RES 2 SULEFATION.
 FT MOD_RES 8 AMIDATION.
 SQ SEQUENCE 8 AA; 1064 MW; DDCAA68378768B5A CRC64;

Query Match 27.9%; Score 12; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
 DB 7 DF 8

RESULT 14
 ID COXG_RAT STANDARD; PRT; 8 AA.
 AC P80430;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Cytochrome c oxidase polypeptide VID (EC 1.9.3.1) (AED) (Fragment).
 GN COX6B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=WISTAR; TISSUE=Liver;
 RX MEDLINE=95324529; PubMed=7601105;
 RA Schaeffer H., Noack H., Halanek W., Brandt U., von Jagow G.;
 RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
 amino-terminal sequences suggest identity of the fetal heart and the
 adult liver isoform.";
 RL Eur. J. Biochem. 230:235-241(1995).
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -1- FUNCTION: THIS PROTEIN MAY BE ONE OF THE HEME-BINDING SUBUNITS OF
 THE OXIDASE.
 CC -1- CATALYTIC ACTIVITY: 4 ferrocytochrome c + O(2) = 4 ferricytochrome
 c + 2 H(2)O.
 KM Oxidoreductase; Mitochondrion.
 FT NON_TER 1
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 1039 MW; 8101E9CAA73AE456 CRC64;

Query Match 27.9%; Score 12; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 DF 3
 DB 5 DF 6

RESULT 15
 ID FAR8_CALVO STANDARD; PRT; 8 AA.
 AC P41863;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Calliphramide 8.
 OS Calliphora vomitoria (Blue blowfly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Preygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Oestroidea; Calliphoridae; Calliphora.
 OX NCBI_TaxID=27454;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Thoracic ganglion;
 RX MEDLINE=92196111; PubMed=1549595;
 RA Duvie H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
 RA Rehfeld J.F., Thorpe A.;
 RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
 neuropeptides (designated calliphramides) from the blowfly
 Calliphora vomitoria.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
 CC -1- SIMILARITY: BELONGS TO THE FAR8 (PMRFAMIDE RELATED PEPTIDE)
 FAMILY.
 DR PIR: H41978; H41978.
 DR Neuropeptide; Amidation.
 FT MOD_RES 8
 FT MOD_RES 8
 SQ SEQUENCE 8 AA; 957 MW; 72D40699CAAA44D8 CRC64;

Query Match 27.9%; Score 12; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DF 3
 DB 4 DF 5

Search completed: July 15, 2002, 13:46:33
 Job time: 192 sec

4.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:43:01 ; Search time 23.8 Seconds
(without alignments)
65.418 Million cell updates/sec

Title: US-09-712-819A-7
Perfect score: 43
Sequence: 1 TDFLTRISS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 648

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	39.5	8	2	P77556
2	17	39.5	8	2	P72221
3	17	39.5	8	4	O15902
4	17	39.5	8	7	O29810
5	16	37.2	7	2	P70804
6	16	37.2	8	4	O90MH9
7	16	37.2	8	12	O91U21
8	16	37.2	8	12	O91U19
9	15	34.9	8	4	O9UD24
10	15	34.9	8	6	O9TT78
11	15	34.9	9	2	O9R635
12	15	34.9	9	11	O9QVH9
13	15	34.9	9	13	P83057
14	14	32.6	9	2	O31363
15	14	32.6	9	8	O94VE1
16	14	32.6	9	8	O94VD8

17	14	32.6	9	15	O85723	O85723 simian sarc
18	13	30.2	8	7	O29801	O29801 homo sapien
19	13	30.2	8	7	O29797	O29797 homo sapien
20	13	30.2	8	7	O29812	O29812 homo sapien
21	13	30.2	9	8	O94VR0	O94VR0 varanus g1g
22	13	30.2	9	11	O92012	O92012 mus musculu
23	13	30.2	5	13	O91LX8	O91LX8 retroperito
24	12	27.9	7	2	P82070	P82070 illoria rub
25	12	27.9	8	2	O34028	O34028 sphingomona
26	12	27.9	8	2	O56246	O56246 thermophil
27	12	27.9	8	2	O47273	O47273 escherichia
28	12	27.9	8	3	P87225	P87225 saccharomyc
29	12	27.9	8	4	O60773	O60773 homo sapien
30	12	27.9	8	5	O3N6M5	O3N6M5 toxoplasma
31	12	27.9	8	8	P92422	P92422 psathyrosta
32	12	27.9	8	8	P92373	P92373 maynaldia v
33	12	27.9	8	8	P93985	P93985 aegilops co
34	12	27.9	8	8	P92404	P92404 lophopyrum
35	12	27.9	8	8	P92426	P92426 pseudoroegn
36	12	27.9	8	8	P93973	P93973 eremopyrum
37	12	27.9	8	8	P93970	P93970 eremopyrum
38	12	27.9	8	8	P92388	P92388 henaradia p
39	12	27.9	8	8	P92428	P92428 peridictyon
40	12	27.9	8	8	P92391	P92391 heteranthel
41	12	27.9	8	8	P92227	P92227 crithopsis
42	12	27.9	8	8	P93963	P93963 psathyrosta
43	12	27.9	8	8	P93961	P93961 psathyrosta
44	12	27.9	8	8	P92215	P92215 amblyopyrum
45	12	27.9	8	8	P93981	P93981 crithodium

ALIGNMENTS

```
RESULT 1
P77556 PRELIMINARY: PRT: 8 AA.
AC P77556:
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
DE TRAY (FRAGMENT).
GN TRAY.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae.
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RC STRAIN=ECOR11.
RA MEDLINE=96400908; PubMed=8807284;
RX Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;
RT "Mosaic structure of plasmids from natural populations of Escherichia
col.,"
RL Genetics 143:1091-1100(1996).
DR EMBL: U50661; AAC44245.1; -
DR EMBL: U50650; AAC44234.1; -
DR EMBL: U50651; AAC44235.1; -
DR EMBL: U50652; AAC44236.1; -
DR EMBL: U50653; AAC44237.1; -
DR EMBL: U50654; AAC44238.1; -
DR EMBL: U50655; AAC44239.1; -
DR EMBL: U50656; AAC44240.1; -
DR EMBL: U50657; AAC44241.1; -
DR EMBL: U50658; AAC44242.1; -
DR EMBL: U50659; AAC44243.1; -
DR EMBL: U50660; AAC44244.1; -
KW Plasmid.
FT NON_TER
SQ SEQUENCE 8 AA; 834 MW; D335A580544735A1 CRC64;
```

Query Match
Best Local Similarity 39.5%; Score 17; DB 2; Length 8;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 TITSS 9
DB 2 SLITSS 7

RESULT 2

P72221 PRELIMINARY; PRT; 8 AA.
AC P72221; 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALGINATE LYASE (FRAGMENT).
GN ALY.
OS Pseudomonas sp. (strain OS-ALG-9).
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=86038;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-OS-ALG-9;
RA Fujiyama K.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-OS-ALG-9;
RA MEDLINE=93329366; PubMed=8336113;
RT Maki H., Mori A., Fujiyama K., Kinoshita S., Yoshida T.;
RT "Cloning, sequence analysis and expression in Escherichia coli of a
RT gene encoding an alginatase from Pseudomonas sp. OS-ALG-9.";
RT J. Gen. Microbiol. 139:987-993(1993).
DR EMBL: D38469; BAA21704.1; -.
KW Lyase.
FT NON_TER
FT NON_TER
SQ SEQUENCE 8 AA: 841 MW; 461DDCC5A5B041BB CRC64;

Query Match
Best Local Similarity 39.5%; Score 17; DB 2; Length 8;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TISS 9
DB 2 TISS 5

RESULT 3
Q15902 PRELIMINARY; PRT; 8 AA.
AC Q15902; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE (CLONE XP7E7A) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.T., Chinault C.A., Baldini A., Lindsey E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrays cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL: L33081; AAA73892.1; -.
FT NON_TER 1 1

FT NON_TER 8 8
SQ SEQUENCE 8 AA: 931 MW; 83D699CAB1B2C9 CRC64;

Query Match
Best Local Similarity 39.5%; Score 17; DB 4; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TDF 3
DB 4 TDF 6

RESULT 4
Q29810 PRELIMINARY; PRT; 8 AA.
AC Q29810; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HLA CLASS II DR-BETA CHAIN (FRAGMENT).
GN HLA-DRB5*15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96175156; PubMed=9110934;
RA Svensson A.C., Selterblad N., Pihlgren U., Raak L., Andersson G.;
RT "Evolutionary relationship between different human major
RT histocompatibility complex HLA-DR haplotypes.";
RL Immunogenetics 43:304-314(1996).
DR EMBL: X88792; CAA61271.1; -.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 8 AA: 909 MW; 6DC6C1A9CA731A8 CRC64;

Query Match
Best Local Similarity 39.5%; Score 17; DB 7; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFT 4
DB 4 DFT 6

RESULT 5
P70804 PRELIMINARY; PRT; 7 AA.
AC P70804; 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALGT PROTEIN (FRAGMENT).
GN ALGT.
OS Acetobacter vinelandii.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Acetobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8830682;
RA Behm B.H.A., Ertesvag H., Valla S.;
RT "A new Acetobacter vinelandii mannuronan C-5-epimerase gene (algC) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL: X87973; CAA61230.1; -.
FT NON_TER 1 1
SQ SEQUENCE 7 AA: 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 37.2% Score 16; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 TISS 9
1:11
DB 2 TVSS 5

RESULT 6
O9UMH9 PRELIMINARY; PRT; 8 AA.
AC O9UMH9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE RHE PROTEIN (FRAGMENT).
GN RHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX MEDLINE=97260406; PubMed=9106526;
RA Matsui G., Cherif-Zahar B., Moura I., Carton J.P.;
RT "Characterization of the recombination hot spot involved in the
RT genomic rearrangement leading to the hybrid D-CE-D gene in the DVI
RT phenotype";
RL Am. J. Hum. Genet. 60:808-817(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BLOOD;
RX MEDLINE=90349591; PubMed=1696722;
RA Cherif-Zahar B., Bloy C., Le Van Kim C., Blanchard D., Bailly P.,
RA Hermand P., Salmon C., Carton J.P., Colin Y.;
RT "Molecular cloning and protein structure of a human blood group Rh
RT polypeptide";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247(1990).
DR EMBL: Z97030; CAB09726.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 1049 MW; C007244691FB5AB1 CRC64;

Query Match 37.2% Score 16; DB 4; Length 8;
Best Local Similarity 28.6%; Pred. No. 5.6e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TDFTLTI 7
1:1:1
DB 1 TDYHML 7

RESULT 7
O9IU21 PRELIMINARY; PRT; 8 AA.
AC O9IU21;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NUCLEAR EXPORT PROTEIN NS2 (FRAGMENT).
GN NS.
OS Influenza A virus (A/Hong Kong/491/97(H5N1)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX NCBI_TaxID=155223;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/HONG KONG/491/97;

RA Shaw M.W., Cooper L.A., Xu X., Thompson W.W., Krauss S.L., Guan Y.,
RA Zhou N.N., Klimov A., Cox N.J., Webster R.G., Lim W., Shortridge K.F.,
RA Subbarao K.;
RT "Avian Influenza A H5N1 and H9N2 viruses bearing a specific
RT constellation of nonglycoprotein genes caused illness in humans";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF256191; AAK49324.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 869 MW; D9D9D5A5A2D1A455 CRC64;

Query Match 37.2% Score 16; DB 12; Length 8;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 TISS 9
1:11
DB 3 TVSS 6

RESULT 8
O9IU19 PRELIMINARY; PRT; 8 AA.
AC O9IU19;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NUCLEAR EXPORT PROTEIN NS2 (FRAGMENT).
GN NS.
OS Influenza A virus (A/Hong Kong/503/97(H5N1)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza virus A and B group; Influenza A viruses; Influenza A virus.
OX NCBI_TaxID=155224;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/HONG KONG/503/97;
RA Shaw M.W., Cooper L.A., Xu X., Thompson W.W., Krauss S.L., Guan Y.,
RA Zhou N.N., Klimov A., Cox N.J., Webster R.G., Lim W., Shortridge K.F.,
RA Subbarao K.;
RT "Avian Influenza A H5N1 and H9N2 viruses bearing a specific
RT constellation of nonglycoprotein genes caused illness in humans";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF256192; AAK49326.1; -.
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 869 MW; D9D9D5A5A2D1A455 CRC64;

Query Match 37.2% Score 16; DB 12; Length 8;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 TISS 9
1:11
DB 3 TVSS 6

RESULT 9
O9UD24 PRELIMINARY; PRT; 8 AA.
AC O9UD24;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RHD PROTEIN (FRAGMENT).
GN RHD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=BLOOD;
RA MEDLINE=97260406; PubMed=9106526;
RA Macassi G., Cherif-Zahar B., Mouru I., Cartton J.P.;
RT "Characterization of the recombination hot spot involved in the
RT genomic rearrangement leading to the hybrid D-CE-D gene in the DVI
RT phenotype.";
RL Am. J. Hum. Genet. 60:808-817(1997).
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=93066356; PubMed=1438298;
RA Le Van Kim C., Mouru I., Cherif-Zahar B., Raynal V., Cherrier C.,
RT "Molecular cloning and primary structure of the human blood group Rh
RT polypeptide.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:10925-10929(1992).
DR EMBL: 297031; CAB09727.1; -.
FT NON_TER 1 1
FT SEQUENCE 8 AA: 1042 MW; D296944691FB5AB1 CRC64;

Query Match 34.9%; Score 15; DB 4; Length 8;
Best Local Similarity 28.6%; Pred. No. 5.6e+05;
Matches 2; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 TDFLT1 7
DB 1 TDYHNM 7

RESULT 10
ID 09TT78 PRELIMINARY; PRT; 8 AA.
AC 09TT78;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE THYMIDYLATE SYNTHASE (FRAGMENT).
GN TS.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21015404; PubMed=11130975;
RA Brouillette J.A., Andrew J.R., Venta P.J.;
RT "Estimate of nucleotide diversity in dogs with a pool-and-sequence
RT method.";
RL Mamm. Genome 11:1079-1086(2000).
DR EMBL: AF202073; AAF20918.1; -.
FT NON_TER 1 1
FT SEQUENCE 8 AA: 899 MW; 6731A1E059CAA867 CRC64;

Query Match 34.9%; Score 15; DB 6; Length 8;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 DFTLTI 7
DB 3 DFIHTL 8

RESULT 11
ID 09R635 PRELIMINARY; PRT; 9 AA.
AC 09R635;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE MAJOR OUTER MEMBRANE PROTEIN VARIABLE DOMAIN IV, MOMP VD IV.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92040090; PubMed=1718870;
RA Peterson E.M., Cheng X., Markoff B.A., Fielder T.J., de la Maza L.M.;
RT "Functional and structural mapping of Chlamydia trachomatis species-
RT specific major outer membrane protein epitopes by use of neutralizing
RT monoclonal antibodies.";
RL Infect. Immun. 59:4147-4153(1991).
SO SEQUENCE 9 AA: 976 MW; 9C61B041B7645361 CRC64;

Query Match 34.9%; Score 15; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 TDFLTIS 8
DB 1 TTMPTIS 8

RESULT 12
ID 09QVH9 PRELIMINARY; PRT; 9 AA.
AC 09QVH9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE SUPERINDUCIBLE PROTEIN 24, SIP24-CYCLOPHILIN HOMOLOG, PEAK C.
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE.
RX MEDLINE=92031730; PubMed=1932334;
RA Davis T.R., Tabatabai L., Bruns K., Hamilton R.T., Nilsen-Hamilton M.;
RT "Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize
RT and secrete a cyclophilin-like protein and beta 2-microglobulin.";
RL Biochim. Biophys. Acta 1095:145-152(1991).
SO SEQUENCE 9 AA: 963 MW; D6D2731A9D6D87D CRC64;

Query Match 34.9%; Score 15; DB 11; Length 9;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTL 5
DB 5 FTL 7

RESULT 13
ID P83057 PRELIMINARY; PRT; 9 AA.
AC P83057;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE BRADYKININ-LIKE PEPTIDE ([VAL,THR3,THR6]BRADYKININ).
OS Bombina variegata (yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archaeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=SKIN SECRETION;
RA Chen T.B., Orr D.F., Bjourson A.J., McClean S., Rao P.F., Shaw C.;
RT "Cloning and post-translational processing of frog skin kininogens.";
RT Submitted (Jul-2001) to the SWISS-PROT data bank.

CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 SQ SEQUENCE 9 AA; 1021 MW; 3687D71A9C861B7 CRC64;

Query Match 34.9%; Score 15; DB 13; Length 9;
 Best Local Similarity 75.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TDT 4
 1 1 1
 Db 3 TGFT 6

RESULT 14
 ID 031363 PRELIMINARY; PRT; 9 AA.
 AC 031363;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE OUTER SURFACE PROTEIN C (FRAGMENT).
 GN OSPC.
 OS Borrelia garinii.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 OX NCB1_TaxID=29519;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-PBI;
 RX MEDLINE=97426044; PubMed=9282748;
 RA Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
 Rosa F.;
 RT "The Borrelia burgdorferi circular plasmid cp26: conservation of
 plasmid structure and targeted inactivation of the ospC gene.";
 RL Mol. Microbiol. 25:361-374(1997).
 DR EMBL; U93699; AAC4553.1; -;
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1019 MW; 4864C1A731A44333 CRC64;

Query Match 32.6%; Score 14; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLT 6
 1 1 1
 Db 5 TLT 7

RESULT 15
 ID 094VE1 PRELIMINARY; PRT; 9 AA.
 AC 094VE1;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
 GN COL.
 OS Varanus mertensi.
 OC Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
 OX NCB1_TaxID=62044;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Ast J.C.;
 RT "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
 RL Cladistics 17:0-0(2001).
 DR EMBL; AF407512; AAL10090.1; -;
 KW Mitochondrion.
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 1154 MW; 9E80C7336411A731 CRC64;

Query Match 32.6%; Score 14; DB 8; Length 9;
 Best Local Similarity 100.0%; Pred. No. 5.6e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLT 6
 1 1 1
 Db 2 TLT 4

Search completed: July 15, 2002, 13:46:16
 Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:39:30 ; Search time 28.73 Seconds
(without alignments)
34.795 Million cell updates/sec

Title: US-09-712-819A-7
Perfect score: 43
Sequence: 1 TDFLTRISS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 102553

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	79.1	7	20	AAV40738
2	34	79.1	7	21	AAAB30076
3	31	72.1	7	20	AAV40737
4	31	72.1	7	21	AAAB30075
5	30	69.8	7	20	AAV40736
6	30	69.8	7	21	AAAB30074
7	25	58.1	8	16	AAW19840
8	23	53.5	8	18	AAW19840
9	22	51.2	9	22	AAW28452
10	22	51.2	9	22	AAW28452
11	22	51.2	9	22	AAW28452

12	21	48.8	7	20	AAV40735	S4 derivative #9,
13	21	48.8	7	20	AAV43399	Peptide CDR-H1-7(Y
14	21	48.8	7	21	AAAB30073	Scaffold protein S
15	21	48.8	9	16	AAW79669	Protein kinase A p
16	21	48.8	9	21	AAV57968	Glycogen synthase
17	21	48.8	9	22	AAW72833	FimA peptide SEQ I
18	20	46.5	6	19	AAW31467	Transcriptional ac
19	20	46.5	8	14	AAW34415	la/SSB epitope 63.
20	20	46.5	8	16	AAW34454	Human TSH receptor
21	20	46.5	8	16	AAW34455	Human TSH receptor
22	20	46.5	8	22	AAU08941	Bioactive cyclic p
23	20	46.5	9	15	AAW67110	Antifertility pept
24	20	46.5	9	15	AAW55716	P. falciparum Pf4
25	20	46.5	9	15	AAW33708	Antigen fragment 2
26	20	46.5	9	21	AAW64453	Catherin-related n
27	20	46.5	9	22	AAW50029	Mutant antimesoth
28	19	44.2	5	9	AAW82200	Example of peptide
29	19	44.2	6	20	AAW06466	Epitope tag. Synt
30	19	44.2	6	21	AAW77708	Epitope tag #4 use
31	19	44.2	6	22	AAW13076	Integrin activatin
32	19	44.2	6	22	AAW51422	AU5 epitope used I
33	19	44.2	6	22	AAW97355	AU5 peptide epitop
34	19	44.2	6	22	AAW59859	Amaranthus viridis
35	19	44.2	7	21	AAW52584	Peptide T for trea
36	19	44.2	8	9	AAW82196	Murine Flt3 ligand
37	19	44.2	8	16	AAW66167	Haemophilus influe
38	19	44.2	8	21	AAW01851	Cadherin-12 cell a
39	19	44.2	8	21	AAW61944	Protocadherin cell
40	19	44.2	8	21	AAW63487	Human CERN2 oncoge
41	19	44.2	9	15	AAW37969	EBNA4 416-424 cyto
42	19	44.2	9	16	AAW78826	Epstein Barr virus
43	19	44.2	9	18	AAW00888	Human neuroendocr
44	19	44.2	9	19	AAW21455	Tyrosinase synthe
45	19	44.2	9	19	AAW77115	

ALIGNMENTS

RESULT 1	
AAV40738	
ID	AAV40738 standard; peptide: 7 AA.
XX	
AC	AAV40738;
DT	01-DEC-1999 (first entry)
DE	S4 derivative #12, beta strand of scaffold protein structure.
KW	Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW	tumour; chemotherapeutic agent.
OS	Synthetic.
XX	
PN	EP947582-A1.
PD	06-OCT-1999.
XX	
PF	31-MAR-1998; 98BP-0870065.
XX	
PR	31-MAR-1998; 98BP-0870065.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX	
DR	WPI: 1999-542958/46.
XX	
PT	New scaffold protein, useful for stabilizing antigens used as vaccines
XX	
PS	
XX	Disclosure: Page 6; 105pp; English.

CC Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10⁸ alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX Sequence 7 AA;

Query Match 79.1%; Score 34; DB 20; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLRIS 8
IIIIII
Db 1 dftlrts 7

RESULT 2

AAB30076 AAB30076 standard; Peptide: 7 AA.

XX AAB30076;

DT 09-FEB-2001 (first entry)

XX Scaffold protein SCA S4 peptide SEQ ID NO: 137.

XX Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
XX SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
XX diabetic retinopathy; atherosclerosis.

OS Synthetic.

PN WO200006070-A1.

PD 12-OCT-2000.

PP 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

XX Scaffold composed of single-chain polypeptide having beta sandwich
XX architecture carrying new and randomized peptide sequences useful as
XX supporting framework and carrying antigen- or receptor binding
XX fragments.
PS Disclosure; Page 15; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.

XX Sequence 7 AA;

Query Match 79.1%; Score 34; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 6.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLRIS 8
IIIIII
Db 1 dftlrts 7

RESULT 3

AAY40737 AAY40737 standard; peptide: 7 AA.

XX AAY40737;

DT 01-DEC-1999 (first entry)

XX S4 derivative #11, beta strand of scaffold protein structure.

XX Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
XX tumour; chemotherapeutic agent.

OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PP 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

XX New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure; Page 6; 105pp; English.

XX Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10⁸ alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will

CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX
SQ Sequence 7 AA;

Query Match 72.1%; Score 31; DB 20; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFTLTIS 8
1:|||||
Db 1 dytlis 7

RESULT 4

AAB30075
ID AAB30075 standard; Peptide: 7 AA.

XX
AC AAB30075;

DT 09-FEB-2001 (first entry)

XX Scaffold protein SCA S4 peptide SEQ ID NO: 136.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis, rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

XX
OS Synthetic.

PM WO200060070-A1.

XX
PD 12-OCT-2000.

PP 01-APR-1999; 99WO-EP02283.

XX
PR 01-APR-1999; 99WO-EP02283.

XX
PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 2000-665002/64.

XX
XX Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments

XX
PS Disclosure: Page 15; 68pp; English.

CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB29930-B29939 were used in the
CC production of the proteins of the invention.

XX
SQ Sequence 7 AA;

Query Match 72.1%; Score 31; DB 21; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFTLTIS 8
1:|||||

Db 1 dytlis 7

RESULT 5

AAY40736
ID AAY40736 standard; peptide: 7 AA.

XX
AC AAY40736;

DT 01-DEC-1999 (first entry)

XX S4 derivative #10, beta strand of scaffold protein structure.

KW Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine;
KW tumour; chemotherapeutic agent.

XX
OS Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

XX
PR 31-MAR-1998; 98EP-0870065.

XX
PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufton S, Hoogenboom H, Sablon E;

DR WPI: 1999-542958/46.

XX
XX New scaffold protein, useful for stabilizing antigens used as vaccines

PT
PT Disclosure: Page 6; 105pp; English.

XX
XX Sequences AAY40727-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAY40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide which forms part of a beta sheet. Peptides
CC (AAY40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10⁴ alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functional equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the
CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

XX
SQ Sequence 7 AA;

Query Match 69.8%; Score 30; DB 20; Length 7;
Best Local Similarity 85.7%; Pred. No. 6.4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFTLTIS 8
1:|||||
Db 1 dftlis 7

```
RESULT 6
AAB30074
ID AAB30074 standard; Peptide: 7 AA.
XX
AC AAB30074;
XX
DT 09-FEB-2001 (first entry)
XX
DE Scaffold protein SCA S4 peptide SEQ ID NO: 135.
XX
KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.
XX
OS Synthetic.
XX
PN WO200060070-A1.
XX
PD 12-OCT-2000.
XX
PF 01-APR-1999; 99WO-EP02283.
XX
PR 01-APR-1999; 99WO-EP02283.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Desmet J, Hufton S, Hoogenboom H, Sablon E;
XX
DR WPI: 2000-665002/64.
XX
PT Scaffold composed of single-chain polypeptide having beta sandwich
PT architecture carrying new and randomized peptide sequences useful as
PT supporting framework and carrying antigen- or receptor binding
PT fragments.
XX
PS Disclosure: Page 15; 68pp; English.
XX
CC The present invention is concerned with producing scaffold proteins
CC based upon the human CTLA-4 SCA domain. These scaffold proteins can be
CC used as a scaffold to bind antigen- or receptor-binding fragments. These
CC can be used in the treatment of diseases such as cancer,
CC atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
CC diabetic retinopathy. Sequences AAB2930-B2933 were used in the
CC production of the proteins of the invention.
XX
SQ Sequence 7 AA:

Query Match          69.8%; Score 30; DB 21; Length 7;
Best Local Similarity 85.7%; Pred. NO. 6.4e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 DFTLTIS 8
   |||:|
Db 1 dftltis 7

RESULT 7
AAW19840
ID AAW19840 standard; Peptide: 8 AA.
XX
AC AAW19840;
XX
DT 26-JAN-1998 (first entry)
XX
DE Chimeric adenovirus coat protein heparin binding motif.
XX
KW Adenovirus; vector; coat protein; gene therapy; gene transfer;
KW human; cancer; autoimmune disease; heart disease; infection;
KW heparin.
XX
OS Synthetic.
```

```
XX
PN WO9720051-A2.
XX
PD 05-JUN-1997.
XX
PF 27-NOV-1996; 96WO-US19150.
XX
PR 21-AUG-1996; 96US-0701124.
PR 28-NOV-1995; 95US-0563368.
PR 21-AUG-1996; 96US-0700846.
XX
PA (GENVE-) GENVEEC INC.
XX
PI Brough DE, Kovesdi I, Wickham TJ;
XX
DR WPI: 1997-310606/28.
XX
PT Adenoviral vectors containing chimeric coat protein - bind and enter
PT cells more efficiently, useful for gene therapy of e.g. cancer,
PT autoimmune diseases, etc.
XX
PS Claim 7; Page 19; 121pp; English.
XX
CC This peptide is used as a universal transfer vector (UTV) sequence
CC or as a spacer sequence in novel chimeric adenovirus coat proteins
CC (CP), especially chimeric fibre proteins. It is modeled after
CC fibronectin and exhibits heparin binding properties. Claimed
CC UTVs/spacers are given in AAW19810-11, AAW19813-25, AAW19827, AAW19829,
CC AAW19831-32 and AAW19834-43). Claimed chimeric CPs differ from the
CC wild-type CP by the introduction of the UTV and/or spacer at or
CC near the C-terminus or in an exposed loop. This imparts on the
CC chimeric CP the ability to bind to and enter cells by means of a
CC novel cell surface binding site. Recombinant vectors comprising
CC the chimeric CP are able to enter cells more efficiently than
CC vectors comprising wild-type CP, especially at lower m.o.i. They
CC are especially useful for gene therapy of e.g. cancers, genetic
CC disorders, pathogenic infections, heart disease or autoimmune
CC diseases.
XX
SQ Sequence 8 AA:

Query Match          58.1%; Score 25; DB 18; Length 8;
Best Local Similarity 62.5%; Pred. NO. 6.4e+05;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFLTIS 8
   |::|||
Db 1 tdfltis 8

RESULT 8
AAR77190
ID AAR77190 standard; Protein: 8 AA.
XX
AC AAR77190;
XX
DT 24-NOV-1995 (first entry)
XX
DE Internal sequence of cytochrome P-450OX.
XX
KW Cytochrome; P-450ox; monooxygenase.
XX
OS Synthetic.
XX
PN WO9516041-A.
XX
PD 15-JUN-1995.
XX
PF 28-NOV-1994; 94WO-EP03938.
XX
PR 08-DEC-1993; 93EP-0810860.
XX
```

PA (CIBA) CIBA GEIGY AG.
 CC (UYRO-) UNIV ROYAL VETERINARY & AGRIC.
 XX
 PI Halkier BA, Koch BM, Lindberg MLER B, Sidsesen O;
 PI Lindberg KOLLER B;
 XX
 DR WPI; 1995-224324/29.
 XX
 PT DNA encoding cytochrome P-450 mono:oxygenase(s) - for modifying
 PT biosynthetic production of cyanogenic glycoside(s) and
 PT glucosinolate(s) in crop plant(s) for increased pest resistance
 PT or improved nutritive value
 XX
 PS Example; Page 27; 73pp; English.
 XX
 CC Cytochrome P-450OX resembles the cytochrome P-450 reported to
 CC convert oximes to nitriles in rat liver microsomes. The mol. wt.
 CC of P-450OX as determined by SDS-PAGE is 51 kD. AA sequences are
 CC obtd. by Edman degradation. The internal polypeptides are obtd.
 CC by trypsin digestion of the purified protein and subsequent sepu.
 CC of peptides using reverse phase HPLC. The N-terminal sequence is
 CC given in AAR77188. A cytochrome with a mol. wt. of 51 kD and this
 CC N-terminal sequence is claimed. Internal peptides are given in
 CC AAR77189-R77192.
 CC
 XX
 SQ Sequence 8 AA;
 XX
 Query Match 53.5%; Score 23; DB 16; Length 8;
 Best Local Similarity 80.0%; Pred. NO. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 DFTLT 6
 III:
 Db 2 dftvt 6
 XX
 RESULT 9
 ABB56069
 ID ABB56069 standard; Peptide; 9 AA.
 XX
 AC ABB56069;
 XX
 DT 15-FEB-2002 (first entry)
 XX
 DE Vascular dementia-associated protein isoform (VPI) 269.
 XX
 DE Vascular dementia; VD; VD-associated protein isoform; VPI; screening;
 KW diagnosis; prognosis; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX WO200169261-A2.
 PN
 XX 20-SEP-2001.
 PD
 XX 14-MAR-2001; 2001WO-GB01106.
 PF
 XX 15-MAR-2000; 2000GB-0006285.
 PR 24-NOV-2000; 2000GB-0028734.
 PR 28-NOV-2000; 2000US-0724391.
 XX
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
 XX
 PI Herath HMAc, Parekh RB, Rohlf C;
 XX
 DR WPI; 2001-557937/62.
 XX
 XX Screening, diagnosis or prognosis of vascular dementia (VD), useful for
 PT determining stage of VD and monitoring the effect of VD therapy;
 PT comprises analysing body fluid by 2-dimensional electrophoresis for
 PT features correlated with VD -
 XX

PS Claim 6; Page 35; 151pp; English.
 XX
 XX The invention relates to screening, diagnosis or prognosis of Vascular
 CC dementia (VD) in a subject comprising analysing body fluid from the
 CC subject by 2-dimensional (2-D) electrophoresis to generate a 2-D array of
 CC features containing at least one chosen feature whose relative abundance
 CC correlates with the presence, absence, stage or severity of VD or
 CC predicts the onset or course of VD, especially detecting in a sample of
 CC cerebrospinal fluid (CSF) from the subject one of 223 VD-associated
 CC protein isoforms (VPIs) (ABB55801-ABB56295) as fully defined in the
 CC specification. Detecting VD-associated features and VPI is useful for the
 CC screening, diagnosis or prognosis of VD, for determining the stage or
 CC severity of VD, for identifying a subject at risk of VD or for
 CC monitoring the effect of therapy administered to a subject having VD.
 CC Nucleic acids encoding a VPI or inhibiting the function of a VPI are
 CC useful for the treatment of VD and for gene therapy.
 CC
 XX
 SQ Sequence 9 AA;
 XX
 Query Match 51.2%; Score 22; DB 22; Length 9;
 Best Local Similarity 80.0%; Pred. NO. 6.4e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 DFTLT 6
 III:
 Db 2 dftvs 6
 XX
 RESULT 10
 AAU28452
 ID AAU28452 standard; Peptide; 9 AA.
 XX
 AC AAU28452;
 XX
 DT 03-JAN-2002 (first entry)
 XX
 DE DPI tryptic digest peptide #49.
 XX
 XX Human; depression associated protein isoform; tryptic digest peptide;
 KW DPI; cerebrospinal fluid; CSF; BAD; bipolar affective disorder;
 KW neuropsychiatric disorder; bipolar mood disorder; neuroleptic;
 KW manic-depressive illness; schizoaffective disorder.
 XX
 OS Homo sapiens.
 OS
 XX WO200162787-A1.
 PN
 XX 30-AUG-2001.
 PD
 XX 23-FEB-2001; 2001WO-GB00786.
 PF
 XX 24-FEB-2000; 2000GB-0004412.
 PR 08-DEC-2000; 2000GB-0030050.
 PR 12-DEC-2000; 2000US-0254830.
 XX
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
 XX
 PI Herath HMAc, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
 XX
 DR WPI; 2001-570626/64.
 XX
 XX Novel nucleic acid encoding a protein associated with bipolar affective
 PT disorder, which is used for diagnosis, prophylaxis and therapy of
 PT neuropsychiatric disorders, such as bipolar affective disorder -
 XX
 PS Disclosure; Page 31; 153pp; English.
 XX
 CC The present invention relates to the identification of depression
 CC associated protein isoforms (DPIs), particularly the tryptic digest
 CC peptides of these proteins. Some of the DPIs (AAU28404-AAU28625)
 CC described are decreased in the cerebrospinal fluid (CSF) of BAD
 CC (bipolar affective disorder) subjects, whilst other DPIs

SQ Sequence 9 AA;

QY	2	DEFLT	6
		:	
Db	2	dfcIs	6

RESULT 11

AAU26098
ID AAU26098 standard; Peptide; 9 AA.

AC AAU26098;

DT 18-DEC-2001 (first entry)

DE Depression-Associated Protein Isoform DPI-9 #5.

KM Human: Bipolar Affective Disorder; BMD: Depression-Associated feature;
KM DP: Depression-Associated protein isoform; DPI: Cerebro-spinal fluid;
KM CSF: antidepressant; antimalaric; nootropic; tranquilliser; neuroleptic;
KM attention deficient disorder; schizoaffective disorder;
KM unipolar affective disorder.

OS Homo sapiens.

PN WO200163294-A2.

PD 30-AUG-2001.

PF 23-FEB-2001; 2001WO-GB00791.

PR 24-FEB-2000; 2000GB-0004412.

PR 12-DEC-2000; 2000US-0254830.

PA (OXFO-) OXFORD GLYSCSCIENCES UK LTD.

PI Herath HMC, Parekh RB, Rohlf C;

DR WPI; 2001-582081/65.

PT Preparation for diagnosing or treating bipolar affected disorder (BAD)
PT or unipolar depression, or for screening for modulators, comprises a
PT BAD-associated protein isoform -
XX
Claim 8; Page 31; 163pp; English.
SS

Claim 8; Page 31; 163pp; English.

The invention relates to a preparation comprising an isolated bipolar affected disorder (BAD)-Associated Protein Isoform (DPIs). The DPIs are used to screen, diagnose or prognosis of BAD or unipolar depression, determine the stage or severity of BAD or unipolar depression, subject at risk of developing BAD or unipolar depression, identify a effect of therapy in a subject. They are also used to screen for or identify agents that interact with a DPI. These agents, antibodies against the DPIs, and nucleic acids encoding the DPIs are used to treat or prevent BAD or unipolar depression. Diseases that can be treated are attention deficient disorder, a schizoaffective disorder, a bipolar or a unipolar affective disorder. The DPIs are used in proteomics. The

CC proteomic approach of using DPs for screening, diagnosis or prognosis of
CC BAD or unipolar depression overcomes the problems of using gene
CC expression analysis, such as not being able to obtain central nervous
CC system (CNS) tissue from a living patient under normal circumstances.
CC The present sequence is a DIP decreased in the CSF (cerebro-spinal
CC fluid) of subjects having BAD.

SQ Sequence 9 AA;

Query Match	51.2%;	Score 22;	DB 22;	Length 9;
Best Local Similarity	80.0%;	Pred. No. 6.4e+05;		
Matches	4;	Conservative	1;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY	2	DETLT	6
	:		
Db	2	dftls	6

RESULT 12

ID AAY40735 standard; peptide; 7 AA.

AC AAY40735;

DT 01-DEC-1999 (flrst entry)

DE S4 derivative #9, beta strand of scaffold protein structure

KM Scaffold protein; beta strand; beta sheet; stabilize antigen; vaccine
tumour; chemotherapeutic agent.

05 Synthetic.

PN EP947582-A1.

PD 06-OCT-1999.

PF 31-MAR-1998; 98EP-0870065.

PR 31-MAR-1998; 98EP-0870065.

PA (INNO-) INNOGENETICS NV.

Desmet J, Hufton S, Hoogenboom H, Sablon E; PI

DR WPI; 1999-542958/46.

PT New scaffold protein, useful for stabilizing antigens used as vaccines

PS Disclosure; page 6; 105pp; English.

CC SequencesAAV40721-Y40748 are functionally equivalent derivatives of the
CC S4 peptide (AAV40607) which forms part of a scaffold protein. S4 is a
CC beta strand peptide together which forms part of a beta sheet. Peptides
CC (AAV40601-Y40609) together form a single-chain scaffold protein which
CC contains at least 1 disulfide bond, contains less than 10% alpha helix
CC and contains at least 6 beta-strands. The scaffold protein is constructed
CC of beta strands S1-S6, and may also include beta strands A1-A3, or any
CC functionally equivalent derivative of these sequences. The beta strands
CC form two beta sheets S1/S4/S3 and S6/S5/S2 with each strand connected to
CC the next by hydrogen bonds, which generate a beta sandwich architecture.
CC If the additional beta strands A1-A3 are included in the structure the
CC scaffold is constructed of two beta sheets, with the structures
CC A1/S1/S4/S3 and S6/S5/S2/A2/A3. The beta strands are connected to each
CC other via amino acid loops, where at least one of the loops binds to a
CC receptor or antigen. The scaffold protein is used to stabilize antigens
CC or whole proteins such as receptors, or their fragments. It may be used
CC to bind two separate molecules. For example, one surface of the scaffold
CC may be bound to a protein which binds to a tumour antigen. This will
CC target the complex to tumour cells. Another surface may be bound to a
CC cytotoxic molecule or an autoimmune antibody which may then kill the

CC tumour cells. Therefore the scaffold protein may be used to target
CC chemotherapeutic agents to specific cells. It may also be used to
CC stabilize individual peptides in a peptide library and may be used in
CC diagnostic techniques, and to stabilize antigens used as vaccines.

CC Sequence 7 AA;

Query Match 48.8%; Score 21; DB 20; Length 7;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTTTIS 8
:::||||
DB 2 ystltis 7

RESULT 13
AAV14399
ID AAV14399 standard; peptide; 7 AA.

XX AAV14399;

DE 17-AUG-1999 (first entry)

XX Peptide CDR-H1-7(Y3F) derived from anti-HCV protease MAb 8D4.

KW Complementarity determining region; CDR; monoclonal antibody; MAb;
KW hepatitis C virus; HCV; protease; binding site.

XX Synthetic.

XX JP11127861-A.

PD 18-MAY-1999.

PF 29-OCT-1997; 97JP-0297451.

PR 29-OCT-1997; 97JP-0297451.

PA (NIHA) JAPAN ENERGY CORP.

DR WPI; 1999-350322/30.

PT Neutralized antibody partial peptide derived from hepatitis C virus

PT - useful for inhibiting Hepatitis C Virus (HCV) serine protease

PS activity

XX Example 1; Page 24; 32pp; Japanese.

CC This sequence corresponds to a peptide (CDR-H1-7; AAV14403) derived from
CC the sequence of the heavy chain variable region complementarity
CC determining region (CDR-1 of the anti-hepatitis C virus (HCV) Ser/Thr
CC protease monoclonal antibody (MAb) 8D4 protein. The peptide has a Tyr
CC to Phe amino acid substitution at position 3 compared to the CDR-H1-7
CC peptide. The invention relates to the use of partial peptides
CC (AAV14348-Y14353) from the MAb 8D4 for inhibiting HCV serine protease
CC activity.

CC Sequence 7 AA;

Query Match 48.8%; Score 21; DB 20; Length 7;
Best Local Similarity 80.0%; Pred. No. 6.4e+05;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TDFTL 5
||| |
DB 1 tdfvl 5

RESULT 14
AAB30073

ID AAB30073 standard; Peptide; 7 AA.

AC AAB30073;

DT 09-FEB-2001 (first entry)

DE Scaffold protein SCA S4 peptide SEQ ID NO: 134.

KW Human; CTLA-4; scaffold protein; antigen-binding; receptor-binding;
KW SCA domain; cancer; thrombosis; osteoporosis; rheumatoid arthritis;
KW diabetic retinopathy; atherosclerosis.

XX Synthetic.

XX WO200060070-A1.

PD 12-OCT-2000.

PF 01-APR-1999; 99WO-EP02283.

PR 01-APR-1999; 99WO-EP02283.

PA (INNO-) INNOGENETICS NV.

PI Desmet J, Hufion S, Hoogenboom H, Sablon E;

DR WPI; 2000-665002/64.

XX Scaffold composed of single-chain polypeptide having beta sandwich
XX architecture carrying new and randomized peptide sequences useful as
XX supporting framework and carrying antigen- or receptor binding
XX fragments -
XX Disclosure; Page 15; 68pp; English.

XX The present invention is concerned with producing scaffold proteins
XX based upon the human CTLA-4 SCA domain. These scaffold proteins can be
XX used as a scaffold to bind antigen- or receptor-binding fragments. These
XX can be used in the treatment of diseases such as cancer,
XX atherosclerosis, thrombosis, osteoporosis, rheumatoid arthritis and
XX diabetic retinopathy. Sequences AAB2930-82939 were used in the
XX production of the proteins of the invention.

XX Sequence 7 AA;

Query Match 48.8%; Score 21; DB 21; Length 7;
Best Local Similarity 66.7%; Pred. No. 6.4e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 FTTTIS 8
:::||||
DB 2 ystltis 7

RESULT 15
AAR79669
ID AAR79669 standard; peptide; 9 AA.

XX AAR79669;

DT 26-FEB-1996 (first entry)

DE Protein kinase A phosphorylation site in glycogen synthase site 2.

KW Peptide library; phosphorylation site; protein kinase; substrate;
KW inhibitor; competitor; cellular response; cell cycle control;
KW immune response; transcriptional activation; cell development.

XX Synthetic.

XX WO9518823-A2.

PD 13-JUL-1995.
 XX
 PF 06-JAN-1995; 95WO-US00147.
 XX
 PR 07-JAN-1994; 94US-0178570.
 XX
 PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
 XX
 PI Cantley LC, Songyang Z;
 XX
 DR WPI; 1995-255036/33.
 XX
 PT Determn.of amino acid sequence of protein kinase phosphorylation
 PT site - by phosphorylation of peptide library and sequencing
 PT phospho:peptide(s) formed, also new substrates and their analogues
 PT for modulating or detecting protein kinase
 XX
 PS Example 5; Page 32; 131pp; English.
 XX
 CC An oriented degenerate peptide library of the amino acid formula
 CC AAR79661 was constructed to isolate the amino acid sequences at the
 CC phosphorylation sites of a protein kinase eg. protein kinase A,
 CC cyclin B/p33(cdc2), src family kinases, etc. Peptides which are
 CC phosphorylated are isolated and their amino acid sequences are compared
 CC to known substrate/inhibitor peptide sequences for that protein kinase.
 CC The peptides AAR79662-73 represent phosphorylation sites for protein
 CC kinase A. This peptide sequence is the phosphorylation site in glycogen
 CC synthase site 2.
 CC The isolated peptides can be used to screen cpds. for effects on the
 CC protein kinase activity, generate antibodies to identify native kinase
 CC substrates, or modulate a variety of cellular responses in which protein
 CC kinases are involved eg. cell cycle control, immune response,
 CC transcriptional activation or cell development.
 XX
 SQ Sequence 9 AA;

Query Match 48.8%; Score 21; DB 16; Length 9;
 Best Local Similarity 66.7%; Pred. No. 6.4e+05;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTSS 9
 11::11
 Db 3 tlvss 8

Search completed: July 15, 2002, 13:42:57
 Job time: 207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:41:15 ; Search time 12.84 Seconds
(without alignments)
17.121 Million cell updates/sec

Title: US-09-712-819A-7
Perfect score: 43
Sequence: 1 TDFTLRTSS 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 60703

Minimum DB seq length: 0
Maximum DB seq length: 9

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	53.5	8	2	US-08-656-177A-14
2	23	53.5	8	4	US-09-256-797-14
3	21	48.8	9	1	US-08-178-570-22
4	21	48.8	9	3	US-08-369-643-22
5	21	48.8	9	5	PCT-US95-00147-22
6	20	46.5	9	2	US-08-934-222-103
7	20	46.5	9	2	US-08-933-402-103
8	20	46.5	9	2	US-09-207-621-103
9	20	46.5	9	2	US-08-532-818-103
10	20	46.5	9	3	US-09-231-797-103
11	20	46.5	9	3	US-08-934-224-103
12	20	46.5	9	3	US-08-933-843-103
13	20	46.5	9	4	US-08-934-223-103
14	20	46.5	9	4	US-09-413-492-103
15	20	46.5	9	4	US-09-187-859-3749
16	19	44.2	6	4	US-09-140-084-4
17	19	44.2	7	3	US-08-916-443A-8
18	19	44.2	7	4	US-08-640-737-38
19	19	44.2	8	1	US-08-276-452A-81
20	19	44.2	8	2	US-08-798-744-81
21	19	44.2	8	4	US-09-187-859-1755
22	19	44.2	8	4	US-09-187-859-2977
23	19	44.2	8	5	PCT-US94-05150-11
24	19	44.2	9	2	US-08-529-190B-67
25	19	44.2	9	3	US-08-159-339A-275
26	19	44.2	9	4	US-08-772-282-21
27	19	44.2	9	4	US-09-187-859-2978

28	19	44.2	9	4	US-09-187-859-3935	Sequence 3935, Ap
29	19	44.2	9	5	PCT-US95-02121-3	Sequence 3, Appl
30	18	41.9	7	4	US-09-187-859-637	Sequence 637, App
31	18	41.9	8	4	US-09-400-208B-18	Sequence 18, Appl
32	18	41.9	8	4	US-09-187-859-638	Sequence 638, App
33	18	41.9	8	4	US-09-187-859-640	Sequence 640, App
34	18	41.9	8	4	US-09-187-859-2507	Sequence 2507, Ap
35	18	41.9	8	6	5496924-22	Patent No. 5496924
36	18	41.9	9	1	US-08-245-853-9	Sequence 9, Appl
37	18	41.9	9	1	US-08-573-675-9	Sequence 9, Appl
38	18	41.9	9	2	US-08-986-234-87	Sequence 87, Appl
39	18	41.9	9	3	US-08-159-339A-910	Sequence 910, App
40	18	41.9	9	3	US-08-159-339A-911	Sequence 911, App
41	18	41.9	9	4	US-09-187-859-641	Sequence 641, App
42	18	41.9	9	4	US-09-187-859-2508	Sequence 2508, Ap
43	18	41.9	9	4	US-09-187-859-3643	Sequence 3643, Ap
44	18	41.9	9	4	US-09-187-859-3855	Sequence 3855, Ap
45	18	41.9	9	6	5496924-36	Patent No. 5496924

ALIGNMENTS

RESULT 1
US-08-656-177A-14
; Sequence 14, Application US/08656177A
; Patent No. 5882851
; GENERAL INFORMATION:
; APPLICANT: Koch, Birgit M.
; APPLICANT: Sibbesen, Ole
; APPLICANT: Halper, Barbara Ann
; APPLICANT: Molter, Birger L.
; TITLE OF INVENTION: Cytochrome P-450 Monooxygenases
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5882851artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,177A
; FILING DATE: 08-AUG-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-19783/R/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; US-08-656-177A-14

Query Match 53.5%; Score 23; DB 2; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFTLT 6
|||:1
Db 2 DFTVT 6

RESULT 2

US-09-256-797-14
; Sequence 14, Application US/09256797
; Patent No. 6133417
; GENERAL INFORMATION:
; APPLICANT: KOCH, Birgit M.
; APPLICANT: Sildesen, Ole
; APPLICANT: Halckier, Barbara Ann
; APPLICANT: Moller, Birger L.
; TITLE OF INVENTION: Cytochrome P-450 Monooxygenases
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6133417artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/256,797
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/656,177
; FILING DATE: 08-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-19783/A/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; US-09-256-797-14

Query Match 53.5%; Score 23; DB 4; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 DFTLT 6
|||:1
Db 2 DFTVT 6

RESULT 3

US-08-178-570-22
; Sequence 22, Application US/08178570
; Patent No. 5532167
; GENERAL INFORMATION:
; APPLICANT: Lewis C. Cantley
; APPLICANT: Zhou Song yang
; TITLE OF INVENTION: Substrate Specificity of Protein Kinases
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/178,570
FILING DATE: JANUARY 7, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptidic
FRAGMENT TYPE: Internal
US-08-178-570-22

Query Match 48.8%; Score 21; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTSS 9
|||:1
Db 3 TLTSS 8

RESULT 4
US-08-369-643-22
; Sequence 22, Application US/08369643A
; Patent No. 6004757
; GENERAL INFORMATION:
; APPLICANT: Cantley, Lewis C.
; APPLICANT: Songyang, Zhou
; TITLE OF INVENTION: Substrate Specificity of Protein Kinases
; FILE REFERENCE: CNS-001CP
; CURRENT APPLICATION NUMBER: US/08/369,643A
; CURRENT FILING DATE: 1995-01-06
; EARLIER APPLICATION NUMBER: US 08/178,570
; EARLIER FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:glycogen
; OTHER INFORMATION: synthase, site 2
US-08-369-643-22

Query Match 48.8%; Score 21; DB 3; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTSS 9
|||:1
Db 3 TLTSS 8

RESULT 5
PCT-US95-00147-22
Sequence 22, Application PC/TUS9500147
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Substrate Specificity of Protein Kinases
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, suite 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00147
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,570
FILING DATE: JANUARY 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Decontl, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBT-004CPPC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-00147-22

Query Match 48.8%; Score 21; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTISS 9
DB 3 TLVSVS 8

RESULT 6
US-08-934-222-103
Sequence 103, Application US/08934222
Patent No. 5928896
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
CONSTRAINING GROUPS WHICH FLANK A PROTEIN-PROTEIN INTERACTION
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,222
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-934-222-103

Query Match 46.5%; Score 20; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTISS 8
DB 3 TLTISS 7

RESULT 7
US-08-933-402-103
Sequence 103, Application US/08933402
Patent No. 5948887
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
CONSTRAINING GROUPS WHICH FLANK A PROTEIN-PROTEIN INTERACTI
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,402
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isaacson, John P.

REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ. ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-933-402-103

Query Match 46.5%; Score 20; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTLS 8
|||||
Db 3 TLTLS 7

RESULT 8
US-09-207-621-103
Sequence 103, Application US/09207621
Patent No. 5952465
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction S
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/207,621
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ. ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-207-621-103

Query Match 46.5%; Score 20; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTLS 8
|||||
Db 3 TLTLS 7

Db 3 TLTLS 7

RESULT 9
US-08-532-818-103
Sequence 103, Application US/08532818
Patent No. 5965698
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interacti
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ. ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-532-818-103

Query Match 46.5%; Score 20; DB 2; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TLTLS 8
|||||
Db 3 TLTLS 7

RESULT 10
US-09-231-797-103
Sequence 103, Application US/09231797
Patent No. 6084066
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
APPLICANT: KINI, R. Manjunatha
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interactio
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington

STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,797
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
APPLICATION NUMBER: PCT/US94/04294
FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-231-797-103

Query Match 46.5%; Score 20; DB 3; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 TLITIS 8
|||||
Db 3 TLITIS 7

RESULT 11
US-08-934-224-103
Sequence 103, Application US/08934224
GENERAL INFORMATION:
PATENT No. 6100044
APPLICANT: EVANS, Herbert J.
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,224
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
APPLICATION NUMBER: PCT/US94/04294

FILING DATE: 21-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-934-224-103

Query Match 46.5%; Score 20; DB 3; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 TLITIS 8
|||||
Db 3 TLITIS 7

RESULT 12
US-08-933-843-103
Sequence 103, Application US/08933843
PATENT No. 611069
GENERAL INFORMATION:
APPLICANT: EVANS, Herbert J.
TITLE OF INVENTION: Polypeptides That Include Conformation-
TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interacti
TITLE OF INVENTION: Site
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20007
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,843
FILING DATE: 19-SEPT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/532,818
FILING DATE: 03-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/143,364
FILING DATE: 29-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/051,741
FILING DATE: 23-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Isacson, John P.
REGISTRATION NUMBER: 33,751
REFERENCE/DOCKET NUMBER: 040433/0148
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-933-843-103

Query Match 46.5%; Score 20; DB 3; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TLITIS 8
|||:|
DB 3 TLITIS 7

RESULT 13
US-08-934-223-103

; Sequence 103, Application US/08934223

; Patent No. 6147189

; GENERAL INFORMATION:

; APPLICANT: EVANS, Herbert J.

; TITLE OF INVENTION: Polypeptides That Include Conformation-

; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interaction

; TITLE OF INVENTION: Site

; NUMBER OF SEQUENCES: 153

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: Suite 500, 3000 K Street NW

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20007

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/934,223

; FILING DATE:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/08/532,818

; FILING DATE: 03-MAY-1996

; APPLICATION NUMBER: PCT/US94/04294

; FILING DATE: 21-APR-1994

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: U.S. 08/143,364

; FILING DATE: 29-OCT-1993

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: U.S. 08/051,741

; FILING DATE: 23-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Isaacson, John P.

; REGISTRATION NUMBER: 33,751

; REFERENCE/DOCKET NUMBER: 040433/0148

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-934-223-103

Query Match 46.5%; Score 20; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TLITIS 8
|||:|
DB 3 TLITIS 7

RESULT 14

US-09-413-492-103

; Sequence 103, Application US/09413492

; Patent No. 6258550

; GENERAL INFORMATION:

; APPLICANT: EVANS, Herbert J.

; TITLE OF INVENTION: Polypeptides That Include Conformation-

; TITLE OF INVENTION: Constraining Groups Which Flank A Protein-Protein Interacti

; TITLE OF INVENTION: Site

; NUMBER OF SEQUENCES: 153

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: Suite 500, 3000 K Street NW

; CITY: Washington

; STATE: DC

; COUNTRY: USA

; ZIP: 20007

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/413,492

; FILING DATE:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/532,818

; FILING DATE: 03-MAY-1996

; APPLICATION NUMBER: PCT/US94/04294

; FILING DATE: 21-APR-1994

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: U.S. 08/143,364

; FILING DATE: 29-OCT-1993

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: U.S. 08/051,741

; FILING DATE: 23-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Isaacson, John P.

; REGISTRATION NUMBER: 33,751

; REFERENCE/DOCKET NUMBER: 040433/0148

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-09-413-492-103

Query Match 46.5%; Score 20; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 TLITIS 8
|||:|
DB 3 TLITIS 7

RESULT 15

US-09-187-859-3749

; Sequence 3749, Application US/09187859A

; Patent No. 6358920

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS

; FILE REFERENCE: 100086.407C1

; CURRENT APPLICATION NUMBER: US/09/187,859A

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 4052

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 3749

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

OTHER INFORMATION: Representative cyclic modulating agent based on
OTHER INFORMATION: cadherin-related neuronal receptor cell adhesion
OTHER INFORMATION: recognition sequence
us-09-187-859-3749

Query Match 46.5%; Score 20; DB 4; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 DFTLRSS 9
:11: 11
Db 1 EFTIDSS 8

Search completed: July 15, 2002, 13:43:16
Job time: 121 sec

•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:45:26 ; Search time 14.1 Seconds
(without alignments)
74.963 Million cell updates/sec

Title: US-09-712-819a-8
Perfect score: 58
Sequence: 1 YGRKKRRQKKK 11

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 1325

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	41.4	10	2 A42089	transcription fact
2	20	34.5	10	2 PQ0785	NADH dehydrogenase
3	18	31.0	11	2 IS4193	Rhesus blood group
4	17	29.3	9	2 PC7076	spectrin alpha cha
5	17	29.3	10	2 A33143	pneumadin - rat
6	17	29.3	11	2 J00395	hypothetical prote
7	16	27.6	5	2 PQ0689	photosystem I 10.4
8	16	27.6	7	2 B33541	hypothetical prote
9	16	27.6	10	2 A60410	beta-neoendorphin
10	15	25.9	7	2 E33932	Ig mu chain D regi
11	15	25.9	9	2 S78762	ribosomal protein
12	15	25.9	10	2 S65387	cytochrome c oxida
13	15	25.9	10	2 IS2645	gene B-50 protein
14	15	25.9	11	2 B43669	hypothetical prote
15	15	24.1	6	2 IS1434	H4 histone - Afric
16	15	24.1	9	2 C36730	hutu protein - Kle
17	14	24.1	10	1 GMRO12	lectin GMR2 alpha
18	14	24.1	10	2 S38305	leucosulfakinin-II
19	14	24.1	10	2 B60656	leucosulfakinin II
20	14	24.1	10	2 S71948	matrix metalloprot
21	14	24.1	10	2 PH0923	T-cell receptor be
22	14	24.1	10	4 SI4943	UGA3 leader peptid
23	14	24.1	11	1 GMROL	leucosulfakinin -
24	14	24.1	11	2 A40693	transgelin - sheep
25	14	24.1	11	2 A44755	20alpha-hydroxyste
26	14	24.1	11	2 A60656	perisulfakinin - A
27	14	24.1	11	2 PT0250	Ig heavy chain CRD
28	14	24.1	11	2 A33571	follicstatin - bovi
29	13	22.4	5	2 B61445	leu-enkephalin - b

30	13	22.4	5	2 A61445	Met-enkephalin - b
31	13	22.4	7	2 A60494	antineoplastic gly
32	13	22.4	6	2 A60139	fatty-acid synthas
33	13	22.4	7	2 A60224	Met-enkephalin-Ar
34	13	22.4	8	2 PQ0701	unidentified 6.5/3
35	13	22.4	8	2 PT0311	Ig heavy chain CRD
36	13	22.4	8	2 PT0323	Ig heavy chain CRD
37	13	22.4	8	2 PH1618	Ig H chain V-D-J r
38	13	22.4	9	2 E28854	fibriopeptide B -
39	13	22.4	9	2 A11497	transaldolase (EC
40	13	22.4	9	2 PT0238	Ig heavy chain CRD
41	13	22.4	9	2 PT0285	Ig heavy chain CRD
42	13	22.4	9	2 PT0324	Ig heavy chain CRD
43	13	22.4	9	2 PH1591	Ig H chain V-D-J r
44	13	22.4	9	2 S36850	Ig heavy chain V r
45	13	22.4	9	2 G41946	T-cell receptor ga

ALIGNMENTS

RESULT 1
A42089
transcription factor I-POU protein, alternative splice form - fruit fly (Drosophila m
C:Species: Drosophila melanogaster
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
C:Accession: A42089
R:Treacy, M.N.; Neilson, L.I.; Turner, E.E.; He, X.; Rosenfeld, M.G.
Cell 68, 491-505, 1992
A>Title: Twin of I-POU: a two amino acid difference in the I-POU homeodomain disting
A:Reference number: A42089; MUID:92154655
A:Accession: A42089
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-10 <TR>
A:Cross-references: GB:S82271; NID:9245517; PID:9245518
A>Note: sequence extracted from NCBI backbone (NCBIN:82267, NCBI:82271, NCBI:82273)
C:Genetics:
A:Gene: FlyBase:Ipou
A:Cross-references: FlyBase:FBgn0004418

Query Match 41.4% Score 24; DB 2; Length 10;
Best Local Similarity 57.1% Pred. No. 6.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
OY 2 GKKRRQ 8
DB 1 GKKRRR 7
RESULT 2
P00785
NADH dehydrogenase (EC 1.6.99.3) 27K chain - fava bean mitochondrion (fragment)
N:Alternate names: complex I, 27K chain; NADH-ubiquinone reductase 27K chain
C:Species: mitochondrion Vicia faba (fava bean)
C>Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 17-Mar-1999
C:Accession: P00785
R:Letterm, S.; Boultry, M.
Plant Physiol. 102, 435-443, 1993
A>Title: Purification and preliminary characterization of mitochondrial complex I (NA
A:Reference number: P00775; MUID:94151437
A:Accession: P00785
A:Molecule type: protein
A:Residues: 1-10 <LE>
C:Comment: Complex I, mitochondrial NADH-ubiquinone reductase, is the first of the
ranging from 5K to 75K.
C:Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone
A:Genetics:
A:Genome: mitochondrion
C:Keywords: electron transfer; mitochondrion; oxidoreductase

```

Query Match          34.5%; Score 20; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 GRKKR 7
   |||
Db 5 GKKRK 10

RESULT 3
154193
Rhesus blood group CcEe protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: 154193
R:Cherif-zaher, B.; Le Van Kim, C.; Rouillac, C.; Raynal, V.; Carron, J.P.; Collin, Y.
Genomics 19, 68-74, 1994
A:Title: Organization of the gene (RHCE) encoding the human blood group RhCcEe antigens
A:Reference number: 154193; MID:94245182
A:Accession: 154193
A>Status: preliminary; translated from GB/EMBL/DDJ
A:Molecule type: DNA
A:Residues: 1-11 <RES>
A:Cross-references: GB:S70456; MID:9546795; PIDN:AA014061.1; PID:94261761
C:Genetics:
A:Gene: GDB:RHCE
A:Cross-references: GDB:229957; OMIM:111700
A:Map position: 1p36.2-1p34

Query Match          31.0%; Score 18; DB 2; Length 11;
Best Local Similarity 57.1%; Pred. No. 6.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 YGRKKR 7
   |||
Db 5 YPRSVRR 11

RESULT 4
PC7076
spectrin alpha chain, non-erythroid - mouse (fragment)
N:Alternate names: fodrin alpha chain
C:Species: Mus musculus (house mouse)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: PC7076
R:Tsuigita, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kano, M.; Matsui, T.; Watanabe, Y.;
Electrophoresis 21, 1853-1871, 2000
A:Title: Proteome analysis of mouse brain: Two-dimensional electrophoresis profiles of
A:Reference number: PC7072
A:Accession: PC7076
A:Molecule type: protein
A:Residues: 1-9 <TSU>
A:Experimental source: strain C57BL/6Cr Slc, male; brain, striatum
C:Keywords: brain

Query Match          29.3%; Score 17; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 RKKR 6
   |||
Db 5 RKKK 8

RESULT 5
A33143
pneumadin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 16-Nov-1990 #sequence_revision 16-Nov-1990 #text_change 18-Aug-2000
C:Accession: A33143
R:Batra, V.K.; Mathur, M.; Mitr, S.A.; Kapoor, R.; Kumar, M.A.

```

```

Regul. Pept. 30, 77-87, 1990
A:Title: Pneumadin: a new lung peptide which triggers antidiuresis.
A:Reference number: A33143; MID:91110910
A:Accession: A33143
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <BAT>
C:Superfamily: unassigned animal peptides

Query Match          29.3%; Score 17; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 8.2e+03;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YGRKK 5
   |||
Db 1 YGEPR 5

RESULT 6
J00395
Hypothetical protein (nodB 3' region) - Azorhizobium caulinodans
N:Alternate names: hypothetical 1.4K protein
C:Species: Azorhizobium caulinodans
A>Note: host Sesbania rostrata
C:Accession: J00395
R:Goethals, K.; Gao, M.; Tomekpe, K.; Van Montagu, M.; Holsters, M.
Mol. Gen. Genet. 219, 289-298, 1989
A:Title: Common nodABC genes in Nod locus 1 of Azorhizobium caulinodans: nucleotide s
A:Reference number: J00393; MID:90136519
A:Accession: J00395
A:Molecule type: DNA
A:Residues: 1-11 <GOE>
A:Cross-references: GB:U18897
A:Experimental source: strain ORS571

Query Match          29.3%; Score 17; DB 2; Length 11;
Best Local Similarity 66.7%; Pred. No. 8.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 6 RROKK 11
   |||
Db 2 RRVDDK 7

RESULT 7
P00689
Photosystem I 10.4K H1 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: P00689
R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyra, M.
Plant Physiol. 102, 1259-1267, 1993
A:Title: Molecular heterogeneity of photosystem I. psal, psalE, psalF, psalH and psalI ar
A:Reference number: P00667; MID:94105345
A:Accession: P00689
A:Molecule type: protein
A:Residues: 1-5 <OBO>
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match          27.6%; Score 16; DB 2; Length 5;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YGRK 4
   |||
Db 2 YGDK 5

RESULT 8

```

B33541
 hypothetical protein (T1 5' region) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Feb-1990 #sequence_revision 21-Feb-1990 #text_change 30-Sep-1993
 C:Accession: B33541
 R:Klemenz, R.; Hoffmann, S.; Wernskjold, A.K.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
 A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to
 A:Reference number: A33541; MUID:8934536
 A:Accession: B33541
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-7 <KLE>
 A:Cross-references: GB:M24843

Query Match 27.6%; Score 16; DB 2; Length 7;
 Best Local Similarity 50.0%; Pred. No. 2.8e+05;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 KRRKK 10
 :|||
 DB 2 RRRSAK 7

RESULT 9
 A60410
 beta-neoendorphin / dynorphin precursor - guinea pig
 N:Alternate names: alpha-neoendorphin; proenkephalin B precursor
 C:Species: Cavia porcellus (guinea pig)
 C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 21-Jan-2000
 C:Accession: A60410
 R:Murphy, R.; Turner, C.A.
 Peptides 11, 65-68, 1990
 A:Title: Isolation and microsequence analysis of guinea pig alpha-neo-endorphin.
 A:Reference number: A60410; MUID:90259864
 A:Accession: A60410
 A:Molecule type: protein
 A:Residues: 1-10 <MUR>
 C:Superfamily: proenkephalin
 C:Keywords: neuropeptide; opioid peptide

Query Match 27.6%; Score 16; DB 2; Length 10;
 Best Local Similarity 40.0%; Pred. No. 1.2e+04;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 YGRRKRKK 10
 :|||
 DB 1 YGGRKKRPK 10

RESULT 10
 E33932
 Ig mu chain D region (E7) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 16-Aug-1996
 C:Accession: E33932
 R:Baccala, R.; Vo Quang, T.; Gilbert, M.; Ternynck, T.; Avrameas, S.
 Proc. Natl. Acad. Sci. U.S.A. 86, 4624-4628, 1989
 A:Title: Two murine natural polyclonal antibodies are encoded by nonmutated germ-line
 A:Reference number: A33932; MUID:89282823
 A:Accession: E33932
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-7 <BAC>
 A:Cross-references: GB:M27106
 C:Keywords: immunoglobulin

Query Match 25.9%; Score 15; DB 2; Length 7;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGR 3
 :|||
 DB 4 YGR 6

RESULT 11
 S78762
 ribosomal protein MRP-S12, mitochondrial - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: S78762
 R:Graack, H.R.
 Submitted to the Protein Sequence Database, July 1999
 A:Reference number: S78760
 A:Accession: S78762
 A:Molecule type: protein
 A:Residues: 1-9 <GRA>
 C:Keywords: mitochondrial
 F:1-9/Product: ribosomal protein MRP-S12 (fragment) #status experimental <MAT>

Query Match 25.9%; Score 15; DB 2; Length 9;
 Best Local Similarity 42.9%; Pred. No. 2.8e+05;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 RKRKRKK 9
 :|||
 DB 2 RKRVRPR 8

RESULT 12
 S63387
 cytochrome-c oxidase (EC 1.9.3.1) chain VTI b, cardiac - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
 C:Accession: S63387; S63386
 R:Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.
 Eur. J. Biochem. 230, 235-241, 1995
 A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t
 A:Reference number: S63372; MUID:95324529
 A:Accession: S63387
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <SCH>
 A:Accession: S63386
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <SC2>
 C:Keywords: cardiac muscle; heart; oxidoreductase

Query Match 25.9%; Score 15; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.7e+04;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OKK 10
 :|||
 DB 3 OKK 5

RESULT 13
 I52645
 gene B-50 protein - rat (fragment)
 C:Species: Rattus sp. (rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 17-Mar-1999
 C:Accession: I52645
 R:Ergen, B.J.; Nieland, H.B.; Rensen-de Ieuw, M.G.; Schotman, P.; Gispén, W.H.; Sc
 Brain Res. Mol. Brain Res. 23, 221-234, 1994
 A:Title: Identification of two promoter regions in the rat B-50/GAP-43 gene.
 A:Reference number: I52645; MUID:94335554
 A:Accession: I52645
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-10 <RES>
A:Cross-references: GB:S71492; NID:g560728
C:Genetics:
A:Gene: B-50

Query Match 25.9%; Score 15; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.7e+04;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 6 RRQK 10
DB 6 RRQK 10

RESULT 14
B43669
hypothetical protein (rhdA 5' region) - *Synechococcus* sp. (fragment)
C:Species: *Synechococcus* sp.
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 30-Sep-1993
C:Accession: B43669
R:Laudenbach, D.E.; Ehrhardt, D.; Green, L.; Grossman, A.
J. Bacteriol. 173, 2751-2760, 1991
A:Title: Isolation and characterization of a sulfur-regulated gene encoding a periplasmic
A:Reference number: A43669; MUID:91210163
A:Accession: B43669
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-11 <LAU>
A:Cross-references: GB:M65244

Query Match 25.9%; Score 15; DB 2; Length 11;
Best Local Similarity 75.0%; Pred. No. 1.8e+04;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 KRQ 8
DB 7 KRQ 10

RESULT 15
I51434
H4 histone - African clawed frog (fragment)
C:Species: *Xenopus laevis* (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: I51434
R:Woodland, H.R.; Warrington, J.R.; Ballantine, J.E.M.; Turner, P.C.
Nucleic Acids Res. 12, 4939-4958, 1984
A:Title: Are there major developmentally regulated H4 gene classes in *Xenopus*?
A:Reference number: I51391; MUID:B4247348
A:Accession: I51434
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-6 <MOO>
A:Cross-references: GB:K02304; NID:g214227; PIDN:AAA49738.1; PID:g555517

Query Match 24.1%; Score 14; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GRKK 5
DB 3 GRKK 6

Search completed: July 15, 2002, 13:47:29
Job time: 123 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:47:11 ; Search time 10.15 Seconds
(without alignments)
41.962 Million cell updates/sec

Title: US-09-712-819A-8
Perfect score: 58
Sequence: 1 YGRRKRRQKKK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 410

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	17	29.3	8	1	UH09_RAT P56575 rattus norv
2	17	29.3	10	1	PNEU_RAT P21996 rattus norv
3	17	29.3	11	1	UH02_YEAST P99013 saccharomyc
4	15	25.9	10	1	COXM_RAT P80431 rattus norv
5	14	24.1	9	1	HUTU_KLEAE P12381 klebsiella
6	14	24.1	9	1	NSKI_SARBU P41492 sarcophaga
7	14	24.1	10	1	LSK2_LEUMA P09039 leucophaea
8	14	24.1	10	1	SPI_HALRO O10997 halocynthia
9	14	24.1	11	1	LSK1_LBOMA P04428 leucophaea
10	14	24.1	11	1	LSKP_PERAM P36885 periplaneta
11	13	22.4	7	1	UN06_PINPS P81675 pinus pinas
12	13	22.4	8	1	AL16_CARMA P81819 carcius ma
13	13	22.4	9	1	F1BB_PAPHA P19343 papio hama
14	13	22.4	9	1	TAL3_PICJA P17441 plichia jadi
15	13	22.4	9	1	TKL1_CALVO P41517 galliphora
16	13	22.4	9	1	TKL1_LOCMI P16223 locusta mig
17	13	22.4	10	1	CU30_LOCMI P11735 locusta mig
18	13	22.4	10	1	GON1_ALIMI P317041 alligator m
19	13	22.4	10	1	GON3_ONCKE P203671 oncorhynch
20	13	22.4	10	1	TKL2_LOCMI P16224 locusta mig
21	13	22.4	10	1	TKL3_LOCMI P30249 locusta mig
22	13	22.4	10	1	TKN1_SCYCA P08608 scyllorhinu
23	13	22.4	10	1	TKNB_RANCA P22689 rana catesb
24	13	22.4	10	1	TKS1_ABDAR P42634 aedes aegypt
25	13	22.4	10	1	TKS2_ABDAR P42635 aedes aegypt
26	13	22.4	11	1	CA31_LITCI P82089 litorea cit
27	13	22.4	11	1	CA32_LITCI P82090 litorea cit
28	13	22.4	11	1	TKN1_OPEIN P82026 uperoleia i
29	13	22.4	11	1	TKN2_OPERU P08612 uperoleia r
30	13	22.4	11	1	TKN2_OPERU P08616 uperoleia r
31	13	22.4	11	1	TKNA_GADMO P01290 gadus morhu
32	13	22.4	11	1	TKNA_HORSE P22688 equus cabal
33	13	22.4	11	1	TKNA_RANCA P22688 rana catesb

34	13	22.4	11	1	TKNA_RANRI P29207 rana ridibu
35	13	22.4	11	1	TKN_PHRU P08615 physalaemus
36	12	20.7	8	1	B44K_PORGI P81886 porphyromon
37	12	20.7	9	1	FARP_CALSI P38495 callinectes
38	12	20.7	10	1	RL16_ACHLA P29221 acholeplasm
39	12	20.7	11	1	CEPL_ACHFU P22790 achetina fu
40	12	20.7	11	1	COR2_PERAM P14496 periplaneta
41	11	19.0	3	1	GRM1_HUMAN P01157 homo sapien
42	11	19.0	5	1	UF01_MOUSE P38639 mus musculu
43	11	19.0	7	1	FAR5_HIRME P42564 hirudo medl
44	11	19.0	9	1	F1BB_ERYPA P19346 erythrocebu
45	11	19.0	9	1	F1BB_MACFU P19345 macaca fusc

ALIGNMENTS

RESULT 1					
ID UH09_RAT	STANDARD:	PRT:	8 AA.		
AC P56575;					
DT 15-DEC-1998 (Rel. 37, Created)					
DT 15-DEC-1998 (Rel. 37, Last sequence update)					
DT 15-DEC-1998 (Rel. 37, Last annotation update)					
DE Unknown protein from 2D-page of heart tissue (Spot P9) (Fragment).					
OS Rattus norvegicus (Rat).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OX NCBI_TaxID=10116;					
RN [1]					
RP STRAIN=WISTAR; TISSUE=Heart;					
RA Li X.-P., Plessner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,					
RA Jungblut P.R.;					
RL Submitted (SEP-1998) to the SWISS-PROT data bank.					
CC -I- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN					
CC PROTEIN IS: 8.9, TTS MW IS: 42 KDa.					
FT NON_TER					
FT SEQUENCE	8 AA: 1029 MW: 950775A6C4140B06 CRC64;				
QY	4 KRRQ 8				
DB	1 QERRQ 5				
RESULT 2					
ID PNEU_RAT	STANDARD:	PRT:	10 AA.		
AC P21996;					
DT 01-AUG-1991 (Rel. 19, Created)					
DT 01-AUG-1991 (Rel. 19, Last sequence update)					
DT 16-OCT-2001 (Rel. 40, Last annotation update)					
DE Pneumadin (PNM).					
OS Rattus norvegicus (Rat).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.					
OX NCBI_TaxID=10116;					
RN [1]					
RP SEQUENCE, AND SYNTHESIS.					
RC TISSUE=Lung;					
RX MEDLINE=91110910; PubMed=2274681;					
RA Batra V.K., Mathur M., Mir S.A., Kapoor R., Kumar M.A.;					
RT Regul. Pept. 30:77-87(1990).					
CC -I- FUNCTION: THIS ANTIDIURETIC PEPTIDE TRIGGERS THE RELEASE OF ADH.					
DR PIR: A33143; A33143.					
KW Amidation.					
FT MOD_RES	10	10	AMIDATION.		

SO SEQUENCE 10 AA; 1048 MW; 641D0DAA723276B CRC64;

Query Match
Best Local Similarity 29.3%; Score 17; DB 1; Length 10;

Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YGRKK 5
1 1 1
Db 1 YGEPK 5

RESULT 3

UXB2_YEAST STANDARD; PRT; 11 AA.

AC P99013;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Unknown protein from 2D-page (Spot 2D-000K2f) (Fragment).
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.

RA STRAIN=X2180-1A;
RA Sanchez J.-C., Golaz O., Schaller D., Morch F., Fruhiger S.,
RA Hughes G.J., Appel R.D., Deshusses J., Hochstrasser D.F.;
RL Submitted (Aug-1995) to the SWISS-PROT data bank.
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.20, ITS MW IS: 9.2 Kda.
DR SWISS-2DPAGE; P99013; YEAST.
FT NON_TER
SQ SEQUENCE 11 AA; 1328 MW; EC38021C0DCB42DA CRC64;

Query Match
Best Local Similarity 29.3%; Score 17; DB 1; Length 11;
Matches 3; Conservative 0; Pred. No. 1.8e+03; Indels 0; Gaps 0;

OY 1 YGRK 4
1 1 1
Db 6 YARK 9

RESULT 4

COXM_RAT STANDARD; PRT; 10 AA.

ID COXM_RAT
AC P80431;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide V11b, mitochondrial (EC 1.9.3.1)
(Fragment).
GN COX2B

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RA STRAIN=MISTAR; TISSUE=Liver;
RA MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform".
RL Eur. J. Biochem. 230:235-241(1995).

CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) -> 4 ferricytochrome

CC c + 2 H(2)O.
KM Oxidoreductase; Mitochondrion.

FT NON_TER
SQ SEQUENCE 10 AA; 1210 MW; CFC70EB771A3326 CRC64;

Query Match
Best Local Similarity 25.9%; Score 15; DB 1; Length 10;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OKK 10
1 1 1
Db 3 OKK 5

RESULT 5

HOTU_KLEAE STANDARD; PRT; 9 AA.

AC P12381;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE Urocannase hydratase (EC 4.2.1.49) (Urocannase)
DE hydrolyase (Fragment).
GN HOTU.
OS Klebsiella aerogenes.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.

ON NCBI_TaxID=28451;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88198018; PubMed=2834335;
RA Nieuwkoop A.J., Baldauf S.A., Hudspeth M.E.S., Bender R.A.;
RT "Bidirectional promoter in the hnt(p) region of the histidine
RT utilization (hnt) operons from Klebsiella aerogenes";
RL J. Bacteriol. 170:2240-2246(1988).

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368611; PubMed=2203754;
RA Schwacha A., Bender R.A.;
RT "Nucleotide sequence of the gene encoding the repressor for the
RT histidine utilization genes of Klebsiella aerogenes";
RL J. Bacteriol. 172:5477-5481(1990).
CC -1- CATALYTIC ACTIVITY: 4,5-dihydro-4-oxo-5-imidazolepropanoate -
CC urocanate + H(2)O.
CC -1- COFACTOR: 1 TIGHTLY BOUND NAD(+) PER CATALYTIC SUBUNIT (BY
CC SIMILARITY).
CC -1- PATHWAY: SECOND STEP IN HISTIDINE DEGRADATION.
CC -1- SIMILARITY: BELONGS TO THE UROCANASE FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: M19665; AAA25078.1; -;
DR EMBL: M34604; AAA25076.1; -;
DR PIR: C36730; C36730.
DR InterPro: IPR000193; Urocannase.
DR PROSITE: PS01233; UROCANASE; PARTIAL.
KW Histidine metabolism; Lyase; NAD.
FT NON_TER
SQ SEQUENCE 9 AA; 1140 MW; 970PC41B5325A6C5 CRC64;

Query Match
Best Local Similarity 24.1%; Score 14; DB 1; Length 9;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 RKKRQ 8

Db 3 QSKYRQ 8

RESULT 6
NSK1_SARBU STANDARD; PRT; 9 AA.
ID NSK1_SARBU
AC P41492;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neosulfakinin-I (NEB-SK-I).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Oestridae; Sarcophagidae; Sarcophaga.
RN NCBI_TaxID=7385;
RN [1]
RP TISSUE-Head;
RC MEDLINE=93083101; PubMed=1360367;
RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;
RT "Isolation and primary structure of two sulfakinin-like peptides from
the fleshly, Neobellieria bullata."
RL Comp. Biochem. Physiol. 103C:135-142(1992).
CC -1- FUNCTION: MYOTROPIC PEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR InterPro: IPR001651; Gastrin.
DR PROSITE: PS00259; GASTRIN; 1.
KW Neuropeptide; Amidation; Sulfation.
FT MOD_RES 4 4 SULFATION (POTENTIAL).
FT MOD_RES 9 9 AMIDATION (POTENTIAL).
SQ SEQUENCE 9 AA; 1187 MW; 8B0A0691E86B5AAA CRC64;

Query Match 24.1%; Score 14; DB 1; Length 9;
Best Local Similarity 40.0%; Pred. No. 1e+05; 2; Indels 0; Gaps 0;
Matches 2; Conservative 1; Mismatches 2;

OY 1 YGRKK 5
|| :
DB 4 YGHMR 8

RESULT 7
LSK2_LEUMA STANDARD; PRT; 10 AA.
ID LSK2_LEUMA
AC P09039;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Leucosulfakinin-II (LSK-II).
OS Leucosphaea maderae (Madeira cockroach), and
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
CC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OX Blattellidae; Blattellidae; Leucosphaea.
RN NCBI_TaxID=6988, 6978;
RN [1]
RP SEQUENCE.
RC SPECIES=L. maderae;
RA MEDLINE=87048769; PubMed=3778455;
RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;
RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with
homology to cholecystokinin and gastrin."
RL Biochem. Biophys. Res. Commun. 140:357-364(1986).
RN [2]
RP SEQUENCE.
RC SPECIES=P. americana; TISSUE=Corpora cardiaca;
RA MEDLINE=90137190; PubMed=2615921;
RA Veenstra J.A.;
RT "Isolation and structure of two gastrin/CKK-like neuropeptides from
the American cockroach homologous to the leucosulfakinins.";

RL Neuropeptides 14:145-149(1989).
CC -1- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
THE COCKROACH HINDGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.
CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
DR PIR: A26335; GMR02.
DR PIR: B60656; B60656.
DR InterPro: IPR001651; Gastrin.
DR PROSITE: PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 5 5 SULFATION (IN L. MADERAE, BUT NOT IN
P. AMERICANA).
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1255 MW; 9B4F5391E86B5AAA CRC64;

Query Match 24.1%; Score 14; DB 1; Length 10;
Best Local Similarity 40.0%; Pred. No. 5e+03; 2; Indels 0; Gaps 0;
Matches 2; Conservative 1; Mismatches 2;

OY 1 YGRKK 5
|| :
DB 5 YGHMR 9

RESULT 8
SPI_HALRO STANDARD; PRT; 10 AA.
ID SPI_HALRO
AC Q10997;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Serine proteinase inhibitor (Fragment).
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
CC Stolidobranchia; Pyridae; Halocynthia.
OX NCBI_TaxID=7729;
OX [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=96321313; PubMed=8759295;
RA Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
RT "Purification and characterization of a 58,000-Da proteinase
inhibitor from the hemolymph of a solitary ascidian, Halocynthia
roretzi."
RL Comp. Biochem. Physiol. 114B:1-9(1996).
CC -1- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
CC -1- SUBUNIT: MONOMER.
DR InterPro: IPR000215; Serpin.
DR PROSITE: PS00284; SERPIN; PARTIAL.
KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1104 MW; 4225C73B1B187AA3 CRC64;

Query Match 24.1%; Score 14; DB 1; Length 10;
Best Local Similarity 42.9%; Pred. No. 5e+03; 2; Indels 0; Gaps 0;
Matches 3; Conservative 2; Mismatches 2;

OY 4 KRRKK 10
|| :
DB 2 KRDGEK 8

RESULT 9
LSK1_LEUMA STANDARD; PRT; 11 AA.
ID LSK1_LEUMA
AC P04428;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Leucosulfakinin-I (LSK-I).

OS Leucophaea maderae (Madelira cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blaberoidea; Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=86315858; PubMed=3749893;
 RA Nechman R.J., Holman G.M., Haddon W.F., Ling N.;
 RT "Leucosulafakinin, a sulfated insect neuropeptide with homology to
 RT gastrin and cholecystokinin."; Science 234:71-73(1986).
 RL
 CC -1- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
 CC THE COCKROACH HINDGUT. INHIBITS MUSCLE CONTRACTION OF HINDGUT.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 CC PIR: A01622; GMROL.
 DR Interpro: IPR001651; Gastrin.
 DR PROSITE: PS00259; GASTRIN; 1.
 KW Hormone; Amidation; Sulfation.
 FT MOD_RES 6 6 SULFATION.
 FT MOD_RES 11 11 AMIDATION.
 FT SEQUENCE 11 AA; 1459 MW; 7E4E0680E86B5AAB CRC64;

Query Match 24.1%; Score 14; DB 1; Length 11;
 Best Local Similarity 40.0%; Pred. No. 5.4e+03;
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 YGRKK 5
 Db 6 YGHMR 10

RESULT 10
 LSKP_PERAM STANDARD; PRT; 11 AA.
 ID LSKP_PERAM
 AC P36885;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Perisulafakinin (Pea-SK-1).
 OS Periplaneta americana (American cockroach).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
 OC Blattodea; Blattidae; Periplaneta.
 OX NCBI_TaxID=6978;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Corpora cardiaca;
 RC MEDLINE=90137190; PubMed=2615921;
 RA Venestra J.A.;
 RT *Isolation and structure of two gastrin/CK-like neuropeptides from
 RT the American cockroach homologous to the leucosulafakins.";
 RL Neuropeptides 14:145-149(1989).
 CC -1- FUNCTION: STIMULATES HINDGUT CONTRACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE GASTRIN/CHOLECYSTOKININ FAMILY.
 CC PIR: A60656; A60656.
 DR Interpro: IPR001651; Gastrin.
 DR PROSITE: PS00259; GASTRIN; 1.
 KW Hormone; Amidation; Sulfation.
 FT MOD_RES 6 6 SULFATION.
 FT MOD_RES 11 11 AMIDATION.
 FT SEQUENCE 11 AA; 1445 MW; 8B4E0680E86B5AAB CRC64;

Query Match 24.1%; Score 14; DB 1; Length 11;
 Best Local Similarity 40.0%; Pred. No. 5.4e+03;
 Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 YGRKK 5
 Db 6 YGHMR 10

RESULT 11
 UN06_PINUS STANDARD; PRT; 7 AA.
 ID UN06_PINUS
 AC P81675;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Unknown protein from 2D-page of needles (N141) (Fragment).
 OS Pinus pinaster (Maritime pine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=71647;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Needle;
 RX MEDLINE=99274088; PubMed=10344291.
 RA Costa P., Pionneau C., Bauw G., Dubos C., Bahran N., Kremer A.,
 RA Frigerio J.-M., Plomion C.;
 RT "Separation and characterization of needle and xylem maritime pine
 RT proteins.";
 RT Electrophoresis 20:1098-1108(1999).
 RL
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 6.6, ITS MW IS: 25 kDa.
 FT NON_TER 1 1
 FT NON_TER 7 7
 FT SEQUENCE 7 AA; 823 MW; 69D76724486B5740 CRC64;

Query Match 22.4%; Score 13; DB 1; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YG 2
 Db 2 YG 3

RESULT 12
 AL16_CARMA STANDARD; PRT; 8 AA.
 ID AL16_CARMA
 AC P81819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Carcinustatin 16.
 OS Carcinus maenas (Common shore crab) (Green crab).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
 OC Eubrachyura; Portunioidea; Portunidae; Carcinus.
 OX NCBI_TaxID=6759;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
 RC MEDLINE=98121193; PubMed=9461295;
 RA Duvé H., Johnsen A.H., Maestri J.-L., Scott A.G., Jayos P.P.,
 RA Thorpe A.;
 RT *Isolation and identification of multiple neuropeptides of the
 RT allotostatin superfamily in the shore crab Carcinus maenas.";
 RL Eur. J. Biochem. 250:727-734(1997).
 CC -1- FUNCTION: MAY ACT AS A NEUROTANSMITTER OR NEUROMODULATOR.
 CC -1- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
 CC Neuropeptide; Amidation; Multigene family.
 KW Neuropeptide; Amidation; Multigene family.
 FT MOD_RES 8 8 AMIDATION.
 FT SEQUENCE 8 AA; 813 MW; 7C28B545AB476878 CRC64;

Query Match 22.4%; Score 13; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YG 2
 Db 1

Db 6 YG 7

RESULT 13

FIBR_PAPHA STANDARD: PRT: 9 AA.

AC P19343;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).

OS FGB.

SN Papio hamadryas (Hamadryas baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecinae; Papio.

OX NCBI_TaxID=9557;

RN [1]

RP SEQUENCE.

RX MEDLINE=84161822; PubMed=6423621.

RA Nakamura S., Takenaka O., Takahashi K.;

RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas, and Theropithecus gelada): their amino acid sequences and evolutionary rates and a molecular phylogeny for the baboons.";

RL J. Biochem. 94:1973-1978(1983).

CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET AGGREGATION.

CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.

CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY THROMBIN, WHICH CLAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.

DR PIR: E28854; E28854.

DR InterPro: IPR002181; Fibrinogen C.

DR PROSITE: PS00514; FIBRIN_AC_DOMAIN; PARTIAL.

KM Blood coagulation; Plasma.

FT PEPTIDE 1 9 FIBRINOPEPTIDE B.

FT NON_TER 9 9

SO SEQUENCE 9 AA; 1057 MW; DDFE71E9C7287B06 CRC64;

Query Match 22.4%; Score 13; DB 1; Length 9;

Best Local Similarity 66.7%; Pred. No. 1e+05;

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 YG 3

Db 7 HGR 9

RESULT 14

TAL3_PICJA STANDARD: PRT: 9 AA.

ID TAL3_PICJA

AC P17441;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Transaldolase III (EC 2.2.1.2) (Fragment).

OS Pichia jadinii (Yeast) (Candida utilis).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Pichia.

OX NCBI_TaxID=4903;

RN [1]

RP SEQUENCE.

RX MEDLINE=75145197; PubMed=1092268;

RA Tsolas O., Sun S.C.;

RT "Isolation of a peptide containing a histidyl-cysteinyl sequence from the active center of transaldolase.";

RL Arch. Biochem. Biophys. 167:525-533(1975).

CC -1- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.

CC -1- CATALYTIC ACTIVITY: Sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate -> D-erythrose 4-phosphate + D-fructose 6-phosphate.

CC -1- PATHWAY: Pentose phosphate pathway; nonoxidative part.

CC -1- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY. SUBFAMILY 1.

DR PIR: A11497; A11497.

DR InterPro: IPR001585; Transaldolase.

DR PROSITE: PS00958; TRANSALDOLASE_2; PARTIAL.

DR PROSITE: PS01054; TRANSALDOLASE_1; PARTIAL.

KM Transferrase; Pentose shunt.

FT NON_TER 1 1

FT MOD_RES 9 9

SO SEQUENCE 9 AA; 1033 MW; 325A31A4EB1E058 CRC64;

Query Match 22.4%; Score 13; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YG 2

Db 1 YG 2

RESULT 15

TKCL_CALVO STANDARD: PRT: 9 AA.

ID TKCL_CALVO

AC P41517;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Callitachykinin I.

OS Calliphora vomitoria (Blue blowfly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Oestroidea; Calliphoridae; Calliphora.

OX NCBI_TaxID=27454;

RN [1]

RP SEQUENCE AND SYNTHESIS.

RX MEDLINE=95075727; PubMed=7984492;

RA Lundquist C.T., Clottens F.L., Holman G.M., Nichols R., Nachman R.J., Naessel D.R.;

RT "Callitachykinin I and II, two novel myotropic peptides isolated from the blowfly, Calliphora vomitoria, that have resemblances to tachykinins.";

RL Peptides 15:761-768(1994).

CC -1- FUNCTION: MYOACTIVE PEPTIDE.

CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.

KM Tachykinin; Neuropeptide; Amidation.

FT MOD_RES 9 9

SO SEQUENCE 9 AA; 981 MW; 2417C8B59CC1B7 CRC64;

Query Match 22.4%; Score 13; DB 1; Length 9;

Best Local Similarity 100.0%; Pred. No. 1e+05;

Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YG 2

Db 6 YG 7

Search completed: July 15, 2002, 13:50:14

Job time: 183 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:47:31 ; Search time 23.84 Seconds
(without alignments)
79.821 Million cell updates/sec

Title: US-09-712-819a-8
Perfect score: 58
Sequence: 1 YGRRKKRRKKK 11

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 1418

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	36.2	9	2	Q93E20
2	20	34.5	7	8	P92421
3	20	34.5	7	8	P92385
4	20	34.5	7	8	P92372
5	20	34.5	7	8	P92403
6	20	34.5	7	8	P92425
7	20	34.5	7	8	P92387
8	20	34.5	7	8	P92427
9	20	34.5	7	8	P92390
10	20	34.5	7	8	P92226
11	20	34.5	7	8	P92214
12	20	34.5	7	8	P92430
13	20	34.5	7	8	P92221
14	20	34.5	7	8	P92442
15	20	34.5	7	8	P92381
16	20	34.5	7	8	P92393

17	20	34.5	7	8	P92218	P92218 australopyr
18	20	34.5	7	8	P92440	P92440 thnopyrum
19	20	34.5	7	8	P92210	P92210 agropyron c
20	20	34.5	8	12	O84271	O84271 human papil
21	19	32.8	9	6	O9TF77	O9TF77 bos taurus
22	19	32.8	11	2	O9EU23	O9EU23 escherichia
23	18	31.0	8	12	O84273	O84273 human papil
24	18	31.0	10	11	O9QVE5	O9QVE5 mus sp. pro
25	18	31.0	10	15	O86324	O86324 rous sarcom
26	18	31.0	10	15	O86325	O86325 rous sarcom
27	18	31.0	10	15	O86326	O86326 rous sarcom
28	18	31.0	11	2	O47345	O47345 escherichia
29	17	29.3	7	12	O9YTR0	O9YTR0 human adeno
30	17	29.3	7	12	O9YTR0	O9YTR0 human adeno
31	17	29.3	7	12	O9YVE3	O9YVE3 human adeno
32	17	29.3	9	2	O45852	O45852 clostridium
33	17	29.3	9	10	O940K4	O940K4 arabidopsis
34	17	29.3	10	4	O9UCU6	O9UCU6 homo sapien
35	17	29.3	11	4	O9Y3G2	O9Y3G2 homo sapien
36	16	27.6	10	2	P83160	P83160 anabaena sp
37	16	27.6	10	2	O51812	O51812 escherichia
38	16	27.6	11	5	O9V7K6	O9V7K6 drosophila
39	16	27.6	11	11	O99N81	O99N81 mus musculu
40	15	25.9	9	11	O99MC3	O99MC3 mus musculu
41	15	25.9	11	2	O9RTU8	O9RTU8 pseudomonas
42	14	24.1	8	6	O9BE93	O9BE93 megaplectra n
43	14	24.1	8	6	O9BE92	O9BE92 turslops tr
44	14	24.1	9	10	O9AXH8	O9AXH8 mesembryant
45	14	24.1	9	12	O82622	O82622 avian infec

ALIGNMENTS

RESULT 1
Q93E20 PRELIMINARY: PRT: 9 AA.
AC Q93E20: 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SCPE (FRAGMENT).
GN SCPE.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O90R;
RX MEDLINE=21424698; PubMed=11532154;
RA Franken C., Hase G., Brandt C., Weber-Heynemann J., Martin S.,
RA Lammler C., Podbielski A., Luticken R., Spellerberg B.;
RT "Horizontal gene transfer and host specificity of beta-haemolytic
RT streptococci: the role of a putative composite transposon containing
RT scpb and lmb.";
RL MOL. Microbiol. 41:925-935(2001).
DR EMBL; AF327852; AAL10713.1; --
FT NON_TER 1
SQ SEQUENCE 9 AA; 1146 MW; 543721AB1326C403 CRC64;

Query Match 36.2%; Score 21; DB 2; Length 9;
Best Local Similarity 37.5%; Pred. No. 5.6e+05;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 KRRRKKK 11
DB 2 KRRRKKK 9

RESULT 2
P92421

ID P92421 PRELIMINARY; PRT: 7 AA.
AC P92421: P92419;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DE 01-JAN-1998 (TREMBlrel. 05, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Psathyrostachys fragilis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Psathyrostachys.
OX NCBI_TaxID=37729;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H4372, AND H917; TISSUE=LEAVES;
RA Petersen G., Seberg O.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z77753; CAB01337.1; -.
DR EMBL: Z77752; CAB01334.1; -.
KW Ribosomal protein; Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA: 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
DB 3 KRRR 6

RESULT 3
P92385 PRELIMINARY; PRT: 7 AA.
ID P92385; P92383;
AC P92385; P92383;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-JAN-1998 (TREMBlrel. 05, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Hordeum maritimum (Seaside barley).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Hordeum.
OX NCBI_TaxID=4519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H299, AND H801; TISSUE=LEAVES;
RA Petersen G., Seberg O.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z77763; CAB01367.1; -.
DR EMBL: Z77762; CAB01364.1; -.
KW Ribosomal protein; Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA: 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
DB 3 KRRR 6

RESULT 4
P92372 PRELIMINARY; PRT: 7 AA.

AC P92372;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Haynaldia villosa.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Haynaldia.
OX NCBI_TaxID=40247;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H5561; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilliceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL: Z77741; CAB01301.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA: 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
DB 3 KRRR 6

RESULT 5
P92403 PRELIMINARY; PRT: 7 AA.
ID P92403;
AC P92403;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Lophopyrum elongatum (Tall wheatgrass) (Argopyrum elongatum).
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Lophopyrum.
OX NCBI_TaxID=4588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H6692; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Trilliceae (Poaceae) based on rpoA
sequence data.";
RL Mol. Phylogenet. Evol. 7:217-230(1997).
DR EMBL: Z77743; CAB01307.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 7 AA: 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
DB 3 KRRR 6

```
RESULT 6
P92425 PRELIMINARY; PRT; 7 AA.
AC P92425;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Pseudoroegneria spicata.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Pseudoroegneria.
CX NCBI_TaxID=4604;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H9082; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data."
RL MOL. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77744; CAB01310.1; -.
KW Chloroplast.
FT NON_TER
SO SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KRRR 7
DB 3 KRRR 6

RESULT 7
P92387 PRELIMINARY; PRT; 7 AA.
AC P92387;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Heterardia persica.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Heterardia.
CX NCBI_TaxID=37678;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H5556; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data."
RL MOL. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77748; CAB01322.1; -.
KW Chloroplast.
FT NON_TER
SO SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KRRR 7
DB 3 KRRR 7
```

```
Db 3 KRRR 6

RESULT 8
P92427 PRELIMINARY; PRT; 7 AA.
AC P92427;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Peridictyon sanctum.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Peridictyon.
CX NCBI_TaxID=37683;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H5575; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data."
RL MOL. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77749; CAB01325.1; -.
KW Chloroplast.
FT NON_TER
SO SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KRRR 7
DB 3 KRRR 6

RESULT 9
P92390 PRELIMINARY; PRT; 7 AA.
AC P92390;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN 11 (FRAGMENT).
GN RPS11.
OS Heterantheium piliferum.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Heterantheium.
CX NCBI_TaxID=37679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H5557; TISSUE=LEAF;
RX MEDLINE=97271648; PubMed=9126564;
RA Petersen G., Seberg O.;
RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
sequence data."
RL MOL. Phylogenet. Evol. 7:217-230(1997).
DR EMBL; Z77750; CAB01328.1; -.
KW Chloroplast.
FT NON_TER
SO SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 4 KRRR 7
||||
DB 3 KRRR 6

RESULT 10
P92226

ID P92226 PRELIMINARY; PRT: 7 AA.

AC P92226; 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

NCBI_TaxID=37674; RIBOSOMAL PROTEIN 11 (FRAGMENT).

OS Chloroplast. Crithopsis deileana.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Trilicaceae; Crithopsis.

NCBI_TaxID=37674; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=37674; RIBOSOMAL PROTEIN 11 (FRAGMENT).

RC STRAIN-H5558; TISSUE-LEAF; MEDLINE=97271648; PubMed=9126564; Petersen G., Seberg O.; Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA

RT sequence data. Mol. Phylogenet. Evol. 7:217-230(1997).

RL EMBL: 277751; CAB01331.1; -. DR Chloroplast.

FT NON-TER 1 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
||||
DB 3 KRRR 6

RESULT 11
P92214

ID P92214 PRELIMINARY; PRT: 7 AA.

AC P92214; 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

NCBI_TaxID=4595; RIBOSOMAL PROTEIN 11 (FRAGMENT).

OS Amphibium multivum.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Trilicaceae; Amblyopium.

NCBI_TaxID=4595; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=4595; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=4595; RIBOSOMAL PROTEIN 11 (FRAGMENT).

RC STRAIN-H5572; TISSUE-LEAF; MEDLINE=97271648; PubMed=9126564; Petersen G., Seberg O.; Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA

RT sequence data. Mol. Phylogenet. Evol. 7:217-230(1997).

RL EMBL: 277756; CAB01346.1; -. DR Chloroplast.

FT NON-TER 1 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
||||
DB 3 KRRR 6

RESULT 12
P92430

ID P92430 PRELIMINARY; PRT: 7 AA.

AC P92430; 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

NCBI_TaxID=37682; RIBOSOMAL PROTEIN 11 (FRAGMENT).

OS Chloroplast. Aegilops tauschii (Aegilops squarrosa).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Trilicaceae; Aegilops.

NCBI_TaxID=37682; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=37682; RIBOSOMAL PROTEIN 11 (FRAGMENT).

RC STRAIN-H6668; TISSUE-LEAF; MEDLINE=97271648; PubMed=9126564; Petersen G., Seberg O.; Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA

RT sequence data. Mol. Phylogenet. Evol. 7:217-230(1997).

RL EMBL: 277758; CAB01352.1; -. DR Chloroplast.

FT NON-TER 1 1
SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;
Best Local Similarity 100.0%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRRR 7
||||
DB 3 KRRR 6

RESULT 13
P92221

ID P92221 PRELIMINARY; PRT: 7 AA.

AC P92221; 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

NCBI_TaxID=15371; RIBOSOMAL PROTEIN 11 (FRAGMENT).

OS Bromus inermis (Smooth brome grass).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Bromus.

NCBI_TaxID=15371; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=15371; RIBOSOMAL PROTEIN 11 (FRAGMENT).

NCBI_TaxID=15371; RIBOSOMAL PROTEIN 11 (FRAGMENT).

RC STRAIN-OSA414; TISSUE-LEAF; MEDLINE=97271648; PubMed=9126564; Petersen G., Seberg O.; Phylogenetic analysis of the Trilicaceae (Poaceae) based on rpoA

RT sequence data. Mol. Phylogenet. Evol. 7:217-230(1997).

RL EMBL: 277759; CAB01355.1; -. DR Chloroplast.

FT NON-TER 1 1

SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;

Best Local Similarity 100.0%; Pred. No. 5.6e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRR 7
| | | |
DB 3 KRR 6

RESULT 14

P92442 PRELIMINARY; PRT; 7 AA.

ID P92442

AC P92442

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE RIBOSOMAL PROTEIN 11 (FRAGMENT).

GN RPS11.

OS Taeniattherum caput-medusae (Medusahae).

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OX NCBI_TaxID=37687;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-H10254; TISSUE-LEAF;

RX MEDLINE=97271648; PubMed=9126564;

RA Petersen G., Seberg O.;

RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA

sequence data.";

RL EMBL; 277760; CAB01358.1; -.

KW Chloroplast.

FT NON_TER

SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;

Best Local Similarity 100.0%; Pred. No. 5.6e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRR 7
| | | |
DB 3 KRR 6

RESULT 15

P92381 PRELIMINARY; PRT; 7 AA.

ID P92381

AC P92381

DT 01-MAY-1997 (TREMBLrel. 03, Created)

DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)

DE RIBOSOMAL PROTEIN 11 (FRAGMENT).

GN RPS11.

OS Hordeum brachyantherum.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OX NCBI_TaxID=52712;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-LEAF;

RX MEDLINE=97271648; PubMed=9126564;

RA Petersen G., Seberg O.;

RT "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA

sequence data.";

RL Mol. Phylogenet. Evol. 7:217-230(1997).

DR EMBL; 277761; CAB01361.1; -.

KW Chloroplast.

FT NON_TER

SQ SEQUENCE 7 AA; 894 MW; 673414033277700 CRC64;

Query Match 34.5%; Score 20; DB 8; Length 7;

Best Local Similarity 100.0%; Pred. No. 5.6e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KRR 7
| | | |
DB 3 KRR 6

Search completed: July 15, 2002, 13:50:45
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:43:41 ; Search time 29.02 Seconds
(without alignments)
42.102 Million cell updates/sec

Title: US-09-712-819A-8
Perfect score: 58
Sequence: 1 YGRKKRRQKKK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 147623

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	89.7	11	22	AAU09906 Interleukin 17 (hi
2	49	84.5	11	19	AAW50263 HIV-1 tat protein.
3	49	84.5	11	20	AAV25075 TAT transductio
4	49	84.5	11	20	AAV05415 Tat peptide. Unid
5	49	84.5	11	21	AAW03932 TAT protein trans
6	49	84.5	11	21	AAW03961 Minimal eleven ami
7	49	84.5	11	21	AAW03961 Beta-catenin deriv
8	49	84.5	11	21	AAW03961 Peptide associated
9	49	84.5	11	21	AAW03961 HIV TAT transduct
10	49	84.5	11	21	AAW03961 HIV tat protein tr
11	49	84.5	11	21	AAW03961 Amino acid sequenc

12	49	84.5	11	21	AAW71015 Human immunodefici
13	49	84.5	11	22	AAU09932 Human immunodefici
14	49	84.5	11	22	AAU099812 HIV-1 tat protein
15	49	84.5	11	22	AAE13064 Protein transducti
16	49	84.5	11	22	AAE12891 Human immunodefici
17	49	84.5	11	22	AAW56573 HIV tat protein fr
18	49	84.5	11	22	AAW50221 HIV-1 tat protein
19	49	84.5	11	22	AAE12605 Human immunodefici
20	49	84.5	11	22	AAE12204 Membrane transport
21	49	84.5	11	22	AAW60006 Internalising pept
22	49	84.5	11	22	AAW82757 HIV TAT protein tr
23	49	84.5	11	22	AAW85847 HIV-1 tat protein
24	49	84.5	11	22	AAW04300 Human immunodefici
25	49	84.5	11	22	AAW05268 Human immunodefici
26	49	84.5	11	22	AAW86863 HIV TAT peptide tr
27	49	84.5	11	22	AAW03418 Human immunodefici
28	49	84.5	11	22	AAW03730 Protein transducti
29	49	84.5	11	22	AAW03815 HIV tat protein tr
30	49	84.5	11	22	AAW02973 Protein transducti
31	49	84.5	11	22	AAW68376 Human Chk1 kinase
32	49	84.5	11	22	AAW70458 Human G2 checkpoint
33	49	84.5	11	22	AAW67673 Transduction domai
34	49	84.5	11	22	AAW73305 HIV-1 TAT protein
35	49	84.5	11	22	AAW69170 HIV tat protein tr
36	49	84.5	11	22	AAW69548 HIV TAT protein tr
37	49	84.5	11	22	AAW70481 HIV TAT protein tr
38	49	84.5	11	22	AAW71756 HIV TAT protein tr
39	49	84.5	11	22	AAW71757 HIV-1 TAT fragment
40	45	77.6	11	22	AAW73627 Tat derived peptid
41	42	72.4	10	21	AAW13003 HIV-1 Tat peptide
42	42	72.4	10	21	AAW78335 N-terminal S-benzy
43	42	72.4	10	22	AAW85052 C-Jun amino termin
44	42	72.4	10	22	AAW85312 HIV-1 TAT peptide.
45	42	72.4	10	22	AAW83193

ALIGNMENTS

RESULT 1	AAU09906	AAU09906 standard; Protein: 11 AA.
ID	AAU09906;	
AC	AAU09906;	
DT	14-FEB-2002 (first entry)	
XX	Interleukin 17 (hIL-17) receptor like protein associated peptide #1.	
DE	Interleukin 17 (hIL-17) receptor like protein associated peptide #1.	
KW	Interleukin 17; hIL-17 receptor like protein; immunomodulatory;	
KW	anti-inflammatory; anti-diabetic; immunosuppressive; anti-microbial;	
KW	hepatic; anabolic; anorectic; anti-alzheimer's; anti-parkinsonian;	
KW	anti-convulsant; anti-asthmatic; dermatological; renal; osteopathic;	
KW	vascular; cytostatic; anti-leukemic; anti-fertility; ophthalmological;	
KW	hepatitis; anorexia; cohexia; neuronal dysfunction; lung disease;	
KW	bone disease; vascular disorder; eye disorder; cancer; human.	
XX	Synthetic.	
OS	WO20016859-A2.	
PN	20-SEP-2001.	
XX	15-MAR-2001; 2001WO-US08678.	
PF	16-MAR-2000; 2000US-189816P.	
XX	28-NOV-2000; 2000US-0724460.	
PR	(AMGE-) AMGEN INC.	
XX	Jing S;	
PA	WPI; 2001-611392/70.	
PI		
XX		
DR		

XX Nucleic acids encoding interleukin 17 receptor like polypeptides,
PT useful for preventing, diagnosing and treating, e.g. leukemia, asthma,
PT diabetes, psoriasis and glaucoma -
XX
PS Disclosure: Page 149; 158pp; English.

CC The invention describes novel nucleic acids encoding interleukin (IL) 17
CC receptor like polypeptides useful as vaccines and in gene therapy. These
CC have immunomodulatory, anti-inflammatory, anti-diabetic, anti-microbial,
CC immunosuppressive, hepatic, anabolic, anorectic, anti-alzheimer's, renal,
CC anti-parkinsonian, anti-convulsant, anti-asthmatic, dermatological,
CC osteopathic, vascular, cytostatic, anti-leukaemic, anti-infectivity and
CC ophthalmological activities. The IL-17 receptor like nucleic acids and
CC proteins may be used to prevent and treat diseases associated with
CC inappropriate IL-17 receptor like polypeptide (IL17rlp) expression. These
CC include, for example immune disorders (e.g. inflammation, diabetes and
CC transplant rejection), infections (e.g. hepatitis and septicemia),
CC weight disorders (e.g. anorexia, cachexia and obesity), neuronal
CC dysfunction (e.g. Alzheimer's disease, Parkinson's disease and epilepsy),
CC lung disorders (e.g. cystic fibrosis, asthma and emphysema), skin disease
CC (e.g. eczema and psoriasis), kidney disease (e.g. glomerulonephritis),
CC bone diseases (e.g. osteoporosis and hypercalcaemia), vascular disorders
CC (e.g. stroke and atherosclerosis, cancers (e.g. leukemia, myeloma and
CC breast cancer), reproductive disorders (e.g. infertility and
CC miscarriage), eye disorders (e.g. glaucoma and retinal neuropathy). The
CC DNA and its complements may also be used as diagnostic probes to detect and
CC quantitate the presence of similar nucleic acids in samples and identify
CC patients needing restorative therapy. The IL17rlp may also be used as
CC antigens in the production of antibodies against the proteins and in
CC assays to identify modulators of expression and activity. The
CC anti-IL17rlp antibodies and antagonists may also be used to down regulate
CC expression and activity.
CC Note: This artificial peptide sequence is given in the sequence listing
CC but is not described in the specification.

XX
XX
SQ Sequence 11 AA:

Query Match 89.7%; Score 52; DB 22; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.042;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGKKRRQKK 11
DB 1 YGKKRRQRR 11

RESULT 2
AAW50263
ID AAW50263 standard; Protein; 11 AA.
XX
XX AAW50263;
XX
XX 20-JUL-1998 (first entry)
XX
XX
XX HIV-1 tat protein.
XX
XX
XX Mouse; BH3 interacting domain death agonist; BID; BCL-2 family;
KW apoptosis; regulation; cell death; inflammation; cancer; arthritis;
KW autoimmune disease; viral infection; lymphoproliferative.
XX
XX Human immunodeficiency virus type 1.
XX
XX
XX W09809980-A1.
XX
XX 12-MAR-1998.
XX
XX 09-SEP-1997; 97WO-US15872.
XX
XX 09-SEP-1996; 96US-0706741.
XX
XX (UNIW) UNIV WASHINGTON.

XX
XX Korsmeyer SJ;
PI
XX
XX WPI; 1998-193546/17.
DR
XX
XX BH3 interacting domain death agonist polypeptide - used for treating
PT decreased apoptotic conditions resulting from inflammation etc.
PT
XX
XX Example 8; Page 85; 118pp; English.

CC The present sequence represents the HIV-1 tat protein which is used in
CC an example of the present invention which describes a BH3 interacting
CC domain death agonist (BID) truncated protein. The BID protein, the
CC DNA encoding it or antisense sequences can be used for preventing or
CC treating a decreased apoptotic state of a cell. The decreased apoptotic
CC state that is treated results from a disease such as cancer, viral
CC infections, lymphoproliferative conditions, arthritis, inflammation and
CC autoimmune diseases. Antibodies against the BID protein can be used for
CC detecting a BID polypeptide in a cell or population of cell. The nucleic
CC acid sequence and the BID protein can also be used for treating
CC immunodeficiency disease (including AIDS), senescence, neurodegenerative
CC disease, ischemic and reperfusion cell death, infertility and
CC wound-healing. Primers derived from the nucleic acid encoding the BID
CC protein can be used for detecting/quantitating the protein and for
CC detecting alterations in the nucleic acid encoding the BID protein.

XX
XX
SQ Sequence 11 AA:

Query Match 84.5%; Score 49; DB 19; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGKKRRQKK 11
DB 1 YGKKRRQRR 11

RESULT 3
AAW25075
ID AAW25075 standard; peptide; 11 AA.
XX
XX AAW25075;
XX
XX 24-AUG-1999 (first entry)
XX
XX
XX TAT transduction domain peptide motif.
XX
XX
XX Anti-pathogen; fusion protein; protein transduction domain; PTD; AZT;
KW cytotoxic domain; suppressor; infection; medicament; ddi; ddC; ddT; 3TC;
KW PTC; DAPD; 1592089; GS92; acyclovir; ganciclovir; peniclovir; interferon;
KW apoptosis; virus; HIV; cytomegalovirus; CMV; herpes simplex virus; HSV-1;
KW hepatitis virus; Kaposi's sarcoma-associated herpes virus; KSHV;
KW herpes virus; yellow fever virus; flavivirus; rhinovirus; plasmoidal;
KW transduction efficiency; cytotoxin.
XX
XX
XX Unidentified.
XX
XX
XX W09929721-A1.
XX
XX 17-JUN-1999.
XX
XX
XX 10-DEC-1998; 98WO-US26358.
XX
XX
XX 20-APR-1998; 98US-0082402.
XX
XX 10-DEC-1997; 97US-0069012.
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Dowdy SF;
XX
XX WPI; 1999-394958/33.
XX

PT	New anti-pathogen systems, particularly for virus and plasmidium infections
XX	Claim 65; Page 34; 123pp; English.
PS	This invention describes a novel anti-pathogen system (APS) comprising a fusion protein constructed from a covalently linked protein transduction domain (PTD) and a cytotoxic domain. The APS can be used for suppressing a pathogen infection in a mammal. The method may further comprise administering a medicament e.g. AZT, ddI, ddc, ddV, 3TC, DAPd, 1592U89, CS92, acyclovir, ganciclovir, penciclovir or an interferon. The APS can also be administered to a mammal in the presence of a pathogen to induce apoptosis in a predetermined population of cells. The products can be used for treating mammals suffering from or susceptible to a viral infection or a disease associated with a virus, e.g. HIV, cytomegalovirus (CMV), herpes simplex virus, e.g. type 1 (HSV-1) hepatitis virus, type C (HCV), Kaposi's sarcoma-associated herpes virus (KSHV or human herpes virus 8), yellow fever virus, flavivirus or rhinovirus, or suffering from or susceptible to plasmoidal infection or a disease associated with a plasmoidal infection, e.g. P. falciparum, P. vivax, P. ovale, or P. malariae. The APS exhibits high transduction efficiency and specifically kills or injures cells infected by one or more pathogens. Formation of the cytotoxin is minimized or eliminated in uninfected cells and in infected cells that keep the pathogen inactive. The APS can be specifically tailored to kill or injure cells infected by one or more pathogen strains. This sequence represents a TAT transduction domain peptide motif which is used in the method of the invention.
CC	Sequence 11 AA;
SC	
SO	
Query Match	84.5%; Score 49; DB 20; Length 11;
Best Local Similarity	72.7%; Pred. No. 0.12;
Matches 8; Conservative	3; Mismatches 0; Indels 0; Gaps 0;
OY	1 YGRRKRROKKK 11 :::
Dd	1 ygrkrrqrerr 11
RESULT 4	
AAAY05415	AAAY05415 standard; peptide; 11 AA.
AC	AAAY05415;
DT	02-JUL-1999 (first entry)
DE	Tat peptide.
KW	BH3 domain; cell death agonist; bcl homology domain; BCL-2 family; apoptosis promoter; cancer cell; virus infected cell; inflammation; autoantibody producing cell; cancer; lymphoproliferative condition; arthritis; autoimmune disease; therapy.
OS	unidentified.
PN	WO9916787-A1.
PD	08-APR-1999.
PF	22-SEP-1998; 98WO-US19765.
PR	07-OCT-1997; 97US-0946039.
PR	26-SEP-1997; 97US-0060133.
PA	(UNITV) UNITV WASHINGTON.
PI	Koismeyer SJ;
DR	WPI; 1999-255058/21.
PT	Bcl homology domain 3 polypeptide

XX Claim 7, Page 67, 104pp: English.

XX This sequence represents a Tat peptide.

CC The invention relates to a bcl homology domain 3 (BH3 domain),

CC derived from a proapoptotic member of the bcl-2 family. The

CC BH3 polypeptide can be used in a method for promoting apoptosis in a

CC target cell, especially where the cell is a cancer cell, a virus infected

CC cell or an autocaltobody producing cell. The BH3 polypeptide can be used

CC in therapeutic compositions for treating disease including cancer, other

CC lymphoproliferative conditions, arthritis, inflammation, and autoimmune

CC diseases, which may result from the down regulation of cell death

CC regulation.

XX

XX Sequence 11 AA:

XX

Query Match	84.5%	Score 49	DB 20	Length 11
Best Local Similarity	72.7%	Pred No. 0.12		
Matches 8	Conservative 3	Mismatches 0	Indels 0	Gaps 0

OY 1 YGRRKKRQKK 11
|||||
cellkkrrrrrr 11

Db 1 YGRRKKRQKK 11

RESULT 5
AAB03932
ID AAB03932 standard; peptide; 11 AA.

XX AAB03932;

AC

XX 26-FEB-2001 (first entry)

DT

XX TAT protein transduction domain (internalisation moiety).

DE

XX Modulating agent; beta-catenin; hair loss; hair growth; skin;

KM exfoliation; Alzheimer's disease; gene transcription;

KW cell differentiation; hearing loss; inner ear; hyperacusis; tinnitus;

XX hair regeneration.

XX

OS Human immunodeficiency virus (HIV).

PN WO200059939-A1.

XX 12-OCT-2000.

PD

XX 04-APR-2000; 2000WO-0509174.

PF

XX 05-APR-1999; 9905-0288373.

PR

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

PA

XX Blaschuk OM, Byers S, Gour BJ;

PI

XX WPI: 2000-679355/66.

DR

XX Modulating agents for inhibiting degradation of cytoplasmic

PT beta-catenin, used for e.g. stimulating hair growth or reducing hair

PT loss; inhibiting development of Alzheimer's disease, comprise

PT internalization moiety and amino acid sequence

XX

XX Claim 4; Page 46; 49pp; English.

XX

XX Modulating agents for inhibiting degradation of cytoplasmic

CC beta catenin are described. The modulating agent comprises an

CC internalisation moiety and one or more of an amino acid sequence

CC SYLPS(PC₄GIHS(PC₄IG, or a peptide analogue or peptidomimetic of

CC the amino acid sequence. The modulating agents are useful for the

CC manufacture of a medicament for stimulating hair growth or reducing

CC hair loss, stimulating skin exfoliation, and inhibiting the

CC development of Alzheimer's disease. They may also be used to

CC increase the beta-catenin level in a cell, to stimulate activation

CC of gene transcription in a cell, and to stimulate cell
CC differentiation. They may further be used to ameliorate hearing loss
CC resulting from a variety of inner ear disorders, such as hyperacusis
CC and tinnitus, through regeneration of hair cells of the inner ear.
XX
XX
SO Sequence 11 AA;

Query Match 84.5%; Score 49; DB 21; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGKRRRQKK 11
|||||
DB 1 ygrkkrrrrr 11

RESULT 6

AAB03961 standard; Peptide: 11 AA.

AC AAB03961;

DT 26-FEB-2001 (first entry)

DE Minimal eleven amino acids present in pTAT.

XX Chimeric protein; fusion protein; FLICE like inhibitor protein;
KW FLIP; Fas; TNF; apoptosis; caspase-8; ligand; T cell; thymocyte;
KW tumour specific antigen; immune response; therapy; prophylaxis;
KW diagnosis; HIV; human immunodeficiency syndrome; AIDS;
KW acquired immune deficiency syndrome.
XX
XX Human immunodeficiency virus.
OS

PN W0200059935-A1.

PD 12-OCT-2000.

PF 05-APR-2000; 2000MO-US09002.

PR 05-APR-1999; 99US-0127867.

PR 06-APR-1999; 99US-0128021.

PA (MAYO-) MAYO FOUND MEDICAL EDUCATION & RES.

PA (PAPA/) PAPA C.

PA (ALGE/) ALGECIRAS-SCHMINICH A.

PI Paya C, Algeciras-schminich A;

DR WPI: 2000-664988/64.

PT Fusion polypeptide useful for inhibiting ligand-induced apoptosis,
PT comprises portion of anti-apoptotic polypeptide linked to a transport
PT group

PS Disclosure: Page 30; 89pp; English.

XX A chimeric group or fusion peptide which comprises a portion of an
CC anti-apoptotic polypeptide which inhibits apoptosis of lymphocytes in
CC combination with a transport group is described. The transport group
CC is capable of transporting the chimeric group or fusion peptide
CC across the cell membrane. The anti-apoptotic polypeptide is FLICE-like
CC inhibitor protein (FLIP) which inhibits Fas and TNF mediated apoptosis
CC by inhibiting binding of Caspase-8 to the Fas receptor complex, thus
CC shutting off the downstream Fas signalling pathway. The chimeric group
CC and fusion peptide are useful for inhibiting ligand-induced apoptosis
CC by bringing them into contact with T cells. The chimeric group is
CC useful for expanding T cells in vitro e.g. T cells specific for
CC particular antigens such as tumour-specific antigen, for enhancing
CC immune response and to inhibit the apoptosis of chronically activated
CC T cells e.g. activated CD4+ T cells in HIV infected patients. The
CC chimeric group is also useful for therapeutic, prophylactic or

CC diagnosis of intracellular delivery of small molecules and
CC macromolecules such as anti-apoptotic polypeptides and nucleic
CC acids encoding such polypeptides. Two primers (AA54297, AA54298) were
CC used to amplify the FLIP cDNA for subcloning into the XhoI-NcoI
CC site of the pHA-TAT vector which contains the N-terminal protein
CC transduction domain from the human immunodeficiency virus tat
CC protein. Tat is a preferred transport moiety.
XX
XX
SO Sequence 11 AA;

Query Match 84.5%; Score 49; DB 21; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGKRRRQKK 11
|||||
DB 1 ygrkkrrrrr 11

RESULT 7

AAB27088 standard; Peptide: 11 AA.

AC AAB27088;

DT 15-FEB-2001 (first entry)

DE Beta-catenin derived internalisation moiety SEQ ID NO: 75.

XX Beta-catenin; cadherin-mediated intercellular adhesion;
KW cell differentiation; modulating agent; hair loss; skin exfoliation;
KW internalisation moiety; flanking sequence; transcription; hearing loss.
XX
XX Human immunodeficiency virus.
OS

PN W0200053632-A1.

PD 14-SEP-2000.

PF 07-MAR-2000; 2000MO-CA00222.

PR 09-MAR-1999; 99US-0265107.

PA (UYMC-) UNIV MCGILL.

PI Blaschuk OW, Gour BJ;

DR WPI: 2000-594308/56.

PT Stimulating beta-catenin mediated gene expression, cellular
PT differentiation and hair growth, involves contacting cells with
PT modulating agent capable of inhibiting interaction between alpha and
PT beta catenin

PS Disclosure: Page 19; 77pp; English.

XX The present invention is concerned with methods of modulating the amount
CC of free beta-catenin in the cell, and methods of stimulating the
CC expression of genes involved in cellular differentiation, the
CC transcription of which is under the control of beta-catenin. The
CC peptides given in AAB27053-B27088, AAB27284-B27300 and AAB27310-B27351
CC can be used as modulating agents which interrupt the interaction between
CC alpha and beta catenin, causing increased levels of the latter and
CC stimulating the activation of beta-catenin mediated transcription. This
CC can be used to stimulate cell differentiation, which can then be used to
CC promote hair growth and skin exfoliation. This latter is particularly
CC useful in the improvement of photodamaged skin and to minimise wrinkles.
CC The modulating peptide can also be used to reduce hearing loss resulting
CC from inner ear disorders such as hyperacusis and tinnitus.
XX

SO Sequence 11 AA;


```
OY      1 YGRRRRRQKK 11
XX      |||
DB      1 Ygrrkrrrrr 11

RESULT 10
AAB09907 standard; peptide; 11 AA.
ID      AAB09907
XX
AC      AAB09907;
XX
DT      06-NOV-2000 (first entry)
XX
DE      HIV tat protein transduction domain.
XX
KW      Targeting signal; MUC-1; immunosuppression; autoimmune disorder;
XX      Immune disorder; inflammatory disorder.
OS      Human immunodeficiency virus.
XX
PN      WO200034468-A2.
XX
PD      15-JUN-2000.
XX
PF      09-DEC-1999; 99WO-US29016.
XX
PR      11-DEC-1998; 98US-0111973.
XX
PA      (BIOM-) BIOMIRA INC.
XX
PI      Agrawal B, Longenecker BM;
XX
DR      WPI: 2000-423418/36.
XX
PT      Use of agent capable of intracellularly inhibiting mucin MUC-1 for
XX      inducing T-cell-based immunosuppression and for treating autoimmune
XX      disorders, transplant rejection and inflammatory disorders -
XX      Disclosure: Page 15; 51pp; English.
XX
CC      The present sequence is the tat protein transduction domain from HIV,
XX      which can be used as a targeting signal. It can be used to internalise
XX      sequences, such as MUC-1 antagonists, within the cell. MUC-1 is a
XX      immunosuppressor, and antagonists act to reduce overactive immune
XX      responses. Thus, the peptide can be used to treat inflammatory disorders
XX      such as rheumatoid arthritis, psoriasis, allergic contact dermatitis and
XX      ankylosing spondylitis, autoimmune disorders including myasthenia gravis,
XX      systemic lupus erythematosus, polyarteritis nodosa, Goodpastures
XX      syndrome, isopathic thrombocytopenic purpura, autoimmune haemolytic
XX      anaemia, Graves' disease, rheumatic fever, pernicious anaemia,
XX      insulin-resistant diabetes mellitus, bullous pemphigoid, pemphigus
XX      vulgaris, viral myocarditis, autoimmune thyroiditis, male infertility,
XX      sarcoidosis, allergic encephalomyelitis, multiple sclerosis, Sjorgens
XX      disease, Reiter's disease, Cellac disease, sympathetic ophthalmia and
XX      primary biliary cirrhosis, immune disorders, graft versus host disease
XX      and transplant rejection.
XX
SQ      Sequence 11 AA:
```

```
AC      AAY93542;
XX
DT      25-SEP-2000 (first entry)
XX
DE      Amino acid sequence of a synthetic protein transduction domain.
XX
KW      Protein transduction system; protein transduction domain;
XX      cytotoxic domain; pathogen infection; retroviral infection;
XX      plasmoidal infection; cancer; prostate cancer.
OS      Synthetic.
XX
PN      WO200034308-A2.
XX
PD      15-JUN-2000.
XX
PF      10-DEC-1999; 99WO-US29289.
XX
PR      10-DEC-1998; 98US-0111701.
XX
PA      (UNIM ) UNIV WASHINGTON.
XX
PI      Dowdy SF;
XX
DR      WPI: 2000-431269/37.
XX
PT      Protein transduction system for treating cancer and pathogenic
XX      infections has a fusion protein comprising a protein transduction
XX      domain covalently linked to a cytotoxic domain -
XX      Example 10; Page 71; 127pp; English.
XX
CC      AAY93542-51 represent synthetic protein transduction domains, which
XX      are used in the protein transduction system of the invention. The
XX      specification describes a protein transduction system, which comprises
XX      a fusion protein. This fusion protein has a covalently linked protein
XX      transduction domain and cytotoxic domain. The system is useful for
XX      treating pathogen infection in mammals, infections such as those
XX      caused by CMV, HSV-1, HCV, KSHV, yellow fever virus, flavivirus or
XX      rhinovirus, retroviral infections such as HIV-1, HIV-2, HTLV-3 and/or
XX      LAV, plasmoidal infections associated with P.faciiparum, P.vivax,
XX      P.ovale, P.malariae. It is also useful for treating cancer, especially
XX      prostate cancer.
XX
SQ      Sequence 11 AA:
```

Query Match 84.5%; Score 49; DB 21; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
OY      1 YGRRRRRQKK 11
XX      |||
DB      1 Ygrrkrrrrr 11

RESULT 12
AAY71015
ID      AAY71015 standard; peptide; 11 AA.
XX
AC      AAY71015;
XX
DT      29-AUG-2000 (first entry)
XX
DE      Human immunodeficiency virus TAT peptide.
XX
KW      Phospholamban; PLB; sarcoplasmic reticulum Ca 2+ ATPase; SERCA2a; HIV;
XX      TAT protein; penetratin; transport peptide; cardiomyocyte; inhibitor;
XX      cargo peptide; contractilin; cardiac contractility; cardiact; treatment;
XX      cardiac disease; heart failure; myocardial dysfunction.
XX
OS      Human immunodeficiency virus.
XX
```

PN WO260025804-A2.
XX 11-MAY-2000.
PD
XX
XX
XX 02-NOV-1999; 99WO-US25692.
XX
XX 02-NOV-1998; 98US-0106718.
PR 27-JUL-1999; 99US-0145883.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Chien K, Dillman W, Minamisawa S, He H, Hoshijima M, Meyer M;
PI Scott C, Wang Y, Silverman GJ;
XX
XX WPI: 2000-365393/31.
DR
XX
XX Treating cardiac diseases, e.g. heart failure or myocardial dysfunction
PT comprises enhancing cardiac contractility by inhibiting interaction
PT between phospholamban and sarcoplasmic reticulum calcium 2+ adenosine
PT triphosphatase -
XX
XX
XX Example 5; Page 54; 56pp; English.
PS
XX The patent discloses a method for the treatment of heart failure, using
CC small peptide complexes and recombinant proteins, that induces
CC phospholamban (PLB) deficiency and inhibits the interaction between PLB
CC and sarcoplasmic reticulum Ca 2+ ATPase (SERCA2a) within cardiomyocytes.
CC The peptide complex comprises of transport peptide like penetratin and
CC cargo peptide selected from mutant PLB, native PLB or antibody against
CC PLB protein (contractilin). Penetratin-PLB peptide functions as a
CC dominant negative inhibitor of PLB-SERCA2a interaction, enhances cardiac
CC contractility and reduces blood pressure. This method is useful for the
CC treatment of cardiac disease e.g. heart failure and myocardial
CC dysfunction. The present sequence is the amino terminal, 11 amino acid
CC protein transduction domain of the denatured human immunodeficiency
CC virus (HIV) Tat protein. It belongs to the penetratin class of peptides,
CC with translocating properties having the ability to carry hydrophilic
CC compounds across the plasma membrane. It is used in the construction of
CC recombinant PLB peptides (e.g., Tat-PLB).
XX
XX Sequence 11 AA;
SQ

Query Match 84.5%; Score 49; DB 21; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGRRKRRQKK 11
DB 1 ygrtkrrgrrr 11

RESULT 13
AAU09932 standard; Peptide; 11 AA.
XX
XX AAU09932;
AC
XX
XX 15-JAN-2002 (first entry)
DT
XX
XX Human immunodeficiency virus (HIV) tat partial peptide sequence.
DE
XX
XX Human; fibroblast growth factor-like; FGF-L; HIV tat protein;
KW protein transduction domain; PDT; human immunodeficiency virus.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX WO200168854-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 13-MAR-2001; 2001WO-US08013.
PF
XX

PR 13-MAR-2000; 2000US-188786P.
XX
XX (AMGE-) AMGEN INC.
PA
XX
XX
XX Jing S, Baas MB;
PI
XX
XX WPI: 2001-596910/67.
DR
XX
XX New fibroblast growth factor-like polypeptide and polynucleotide for
PT diagnosis, prevention and treatment of diseases, disorders or
PT conditions involving the central nervous system, teeth, heart, liver or
PT adipose tissue -
XX
XX Disclosure; Page 62; 116pp; English.
PS
XX
XX The present invention relates to new isolated fibroblast growth
CC factor-like (FGF-L) nucleic acid molecules and polypeptides. The FGF-L
CC polypeptide is useful for determining whether a compound inhibits FGF-L
CC polypeptide, and also for identifying a compound that binds to the
CC polypeptide. The FGF-L polypeptides of the invention are useful for
CC treating, preventing or ameliorating a medical condition or an FGF-L
CC polypeptide-related disease, condition or disorder such as wound healing
CC disorders, ulcers, gut disorders, lung disorders, liver disorders such as
CC hepatitis and diabetes. The invention is also useful for diagnosing a
CC pathological condition or susceptibility to a pathological condition in
CC a subject and is useful for modulating levels of FGF-L in an animal.
CC Other uses are detecting or quantifying the amount of FGF-L polypeptide
CC and for identifying or developing novel agonists and antagonists of the
CC FGF-L polypeptide signalling pathway which are useful for treating one
CC or more diseases or disorders, and also as an immunogen for producing
CC antibodies for in vivo imaging. The present sequence represents an 11
CC amino acid sequence from the human immunodeficiency virus (HIV) tat
CC protein that is termed the protein transduction domain or Tat PDT.
XX
XX Sequence 11 AA;
SQ

Query Match 84.5%; Score 49; DB 22; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.12;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 YGRRKRRQKK 11
DB 1 ygrtkrrgrrr 11

RESULT 14
AAU09812 standard; Protein; 11 AA.
XX
XX AAU09812;
AC
XX
XX 27-FEB-2002 (first entry)
DT
XX
XX HIV-1 tat protein transduction domain (TAT PDT).
DE
XX
XX fibroblast growth factor receptor-like protein; FGF-R; anorectic;
KW haemostatic; osteopathic; cytostatic; nephrotoxic; antidiabetic;
KW immunomodulator; antiinflammatory; haematopoietic disorder; osteoporosis;
KW osteogenesis imperfecta; Paget's disease; periodontal disease; cancer;
KW hypercalcaemia; acute glomerulonephritis; chronic glomerulonephritis;
KW diabetes; obesity; cachexia; transgenic animal; gene therapy.
XX
XX human immunodeficiency virus; HIV-1; tat protein.
OS
XX
XX Human immunodeficiency virus type 1.
PN
XX
XX WO200170977-A2.
PD
XX
XX 27-SEP-2001.
PF
XX
XX 22-MAR-2001; 2001WO-US09073.
PR 22-MAR-2000; 2000US-191379P.
XX

XX	(AMGE-)	AMGEN INC.
PA	(SARI/)	SARIS C M.
PA	(MUSX/)	MU S X.
PA	(XIAM/)	XIA M.
PA	(BOON/)	BOONE T C.
PA	(COVE/)	COVER T.
PI	Saris CM,	Mu SX, Xia M, Boone TC, Covey T;
DR	WPI:	2001-626128/72.
XX		
PT	Novel nucleic acid encoding fibroblast growth factor receptor-like	
PT	polypeptides, useful for treating hematopoietic disorder, osteoporosis,	
PT	Paget's disease, glomerulonephritis, cancer, diabetes, obesity and	
PT	cachexia	-
PS	Disclosure; Page 59; 163pp;	English.
XX		
CC	The invention relates to a novel isolated fibroblast growth factor	
CC	receptor-like (FGFR-L) polypeptide (I). (I) and the nucleic acid (II)	
CC	encoding (I) are useful for treating, preventing or ameliorating	
CC	a medical condition including haematopoietic disorder, osteoporosis,	
CC	osteogenesis imperfecta, Paget's disease, periodontal disease,	
CC	hypercalcaemia, acute glomerulonephritis, chronic glomerulonephritis,	
CC	cancer, diabetes, obesity and cachexia. (I) is also useful for	
CC	identifying a compound which binds to FGFR-L polypeptide, by contacting	
CC	(I) with a compound, determining the extent of binding of the FGFR-L	
CC	polypeptide to the compound, and determining the activity of the	
CC	polypeptide when bound to the compound. (II) is useful for modulating	
CC	levels of a polypeptide in an animal. A transgenic animal comprising (I)	
CC	is useful for determining whether a compound inhibits FGFR-L polypeptide	
CC	activity or FGFR-L polypeptide production, by exposing the transgenic	
CC	animal to the compound and measuring FGFR-L polypeptide or production in	
CC	the mammal. (II) is useful for mapping the locations of FGFR-L gene and	
CC	related genes on chromosomes, as hybridisation probes in diagnostic	
CC	assays to test for the presence of an FGFR-L nucleic acid molecule in	
CC	mammalian tissue or bodily fluid samples, in gene therapy, and as tools	
CC	for isolating corresponding FGFR-L polypeptide genes. (I) is useful as	
CC	immunogen, and for cloning FGFR-L polypeptide ligands using an	
CC	expression cloning strategy. The present sequence represents the	
CC	amino acid sequence of human immunodeficiency virus (HIV-1) tat protein	
CC	transduction domain (TAT PDF) as described in the invention.	
XX		
SQ	Sequence	11 AA;
	Query Match	84.5%; Score 49; DB 22; Length 11;
	Best Local Similarity	72.7%; Pred. No. 0.12;
	Matches	8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
OY	1 YGRRKKRQKK 11	
b	1 ygrfkkrrqgrr 11	
	RESULT 15	
ID	AAE13064 standard; peptide; 11 AA.	
XX	AAE13064:	
AC		
XX		
DT	28-JAN-2002 (first entry)	
XX		
DE	Protein transduction domain of HIV tat protein.	
KW	Human; CD20/19F-receptor like protein; immunoglobulin E; agp-96614-a1;	
KW	agp-65406-a1; cancer; abnormal cell proliferation; autoimmune disease;	
KW	ovarian cancer; brain cancer; arteriosclerosis; vascular restenosis;	
KW	rheumatoid arthritis; multiple sclerosis; allergy; dermatitis; asthma;	
KW	reproductive disease; diabetes; transplant rejection; endometriosis;	
KW	infertility; gene therapy; protein transduction domain; HIV;	
KW	human immunodeficiency virus; TAT PDI.	

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:45:51 ; Search time 12.83 Seconds
(without alignments)
20.942 Million cell updates/sec

Title: US-09-712-819A-8
Perfect score: 58
Sequence: 1 YGRRRRRQRRK 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 76679

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	84.5	11	2	US-08-706-741B-54
2	49	84.5	11	2	US-08-924-695A-54
3	49	84.5	11	4	US-09-208-966-2
4	49	84.5	11	4	US-09-298-089-37
5	42	72.4	10	4	US-09-336-093-7
6	42	72.4	11	4	US-09-208-966-16
7	42	72.4	11	4	US-09-208-966-52
8	37	63.8	11	1	US-07-694-983-15
9	36	62.1	7	5	PCT-US95-06077-17
10	36	62.1	9	1	US-08-218-608-3
11	36	62.1	9	1	US-08-357-056-1
12	36	62.1	9	1	US-08-475-583-1
13	36	62.1	9	2	US-08-442-461D-19
14	36	62.1	9	2	US-08-893-853-4
15	36	62.1	9	4	US-08-564-164A-10
16	36	62.1	9	4	US-09-113-921-4
17	36	62.1	9	4	US-09-083-259-1
18	36	62.1	10	1	US-08-071-811A-10
19	36	62.1	10	2	US-08-442-461D-28
20	36	62.1	10	4	US-09-044-411-10
21	36	62.1	11	4	US-09-044-411-1
22	36	62.1	11	4	US-09-044-411-2
23	36	62.1	11	4	US-09-044-411-3
24	36	62.1	11	4	US-09-044-411-4
25	36	62.1	11	4	US-09-044-411-5
26	34	58.6	11	4	US-09-208-966-3
27	33	56.9	8	2	US-08-701-124-1

28	33	56.9	8	3	US-09-130-225-1	Sequence 1, Appl
29	33	56.9	11	2	US-09-455-061-1	Sequence 1, Appl
30	33	56.9	11	4	US-08-701-124-19	Sequence 19, Appl
31	33	56.9	11	3	US-09-130-225-19	Sequence 19, Appl
32	33	56.9	11	4	US-09-208-966-7	Sequence 7, Appl
33	33	56.9	11	4	US-09-455-061-19	Sequence 19, Appl
34	32	55.2	9	1	US-08-332-518-2	Sequence 2, Appl
35	32	55.2	9	1	US-08-332-518-7	Sequence 7, Appl
36	32	55.2	9	2	US-08-378-709-3	Sequence 3, Appl
37	32	55.2	9	2	US-08-378-709-8	Sequence 8, Appl
38	32	55.2	9	4	US-09-083-259-2	Sequence 2, Appl
39	32	55.2	10	1	US-08-097-830E-1	Sequence 1, Appl
40	32	55.2	10	2	US-08-456-112B-1	Sequence 1, Appl
41	32	55.2	11	4	US-09-044-411-6	Sequence 6, Appl
42	32	55.2	11	4	US-08-584-043A-43	Sequence 43, Appl
43	32	55.2	11	4	US-08-584-043A-99	Sequence 99, Appl
44	31	53.4	9	1	US-08-332-518-3	Sequence 3, Appl
45	31	53.4	9	1	US-08-332-518-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-706-741B-54
; Sequence 54, Application US/08706741B
; Patent No. 5955593
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63146
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,741B
; FILING DATE: 09-SEP-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-706-741B-54

Query Match 84.5%; Score 49; DB 2; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.067;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Y 1 YGRRRRRQRRK 11
DB 1 YGRRRRRQRRR 11

```
RESULT 2
US-08-924-695A-54
; Sequence 54, Application US/08924695A
; Patent No. 5998583
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, STANLEY J.
; TITLE OF INVENTION: BH3 INTERACTING DOMAIN DEATH AGONIST
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,695A
; FILING DATE: 09-SEP-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 971798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-924-695A-54

Query Match      84.5%; Score 49; DB 2; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.067;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGKKRRQKK 11
| | | | | | | | | | | | |
Db 1 YGKKRRQRR 11

RESULT 3
US-09-208-966-2
; Sequence 2, Application US/09208966
; Patent No. 622135
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/208,966
; CURRENT FILING DATE: 1998-12-10
; EARLIER APPLICATION NUMBER: 60/082,402
; EARLIER FILING DATE: 1998-04-20
; EARLIER APPLICATION NUMBER: 60/069,012
; EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 11
; TYPE: PRT
; ORGANISM: human
US-09-208-966-2
```

```
Query Match      84.5%; Score 49; DB 4; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.067;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGKKRRQKK 11
| | | | | | | | | | | | |
Db 1 YGKKRRQRR 11

RESULT 4
US-09-296-089-37
; Sequence 37, Application US/09296089
; Patent No. 6303576
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Byers, Stephen
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: BETA-CATENIN MEDIATED GENE EXPRESSION
; FILE REFERENCE: 100086.411
; CURRENT APPLICATION NUMBER: US/09/296,089
; CURRENT FILING DATE: 1999-04-21
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 37
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus
US-09-296-089-37
```

```
Query Match      84.5%; Score 49; DB 4; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.067;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 YGKKRRQKK 11
| | | | | | | | | | | | |
Db 1 YGKKRRQRR 11

RESULT 5
US-09-336-093-7
; Sequence 7, Application US/09336093A
; Patent No. 6348185
; GENERAL INFORMATION:
; APPLICANT: Washington University School of Medicine
; TITLE OF INVENTION: MEMBRANE-PERMEANT PEPTIDE COMPLEXES FOR MEDICAL
; TITLE OF INVENTION: IMAGING, DIAGNOSTICS, AND PHARMACEUTICAL THERAPY
; FILE REFERENCE: WSHU 2001
; CURRENT APPLICATION NUMBER: US/09/336,093A
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TAT protein
US-09-336-093-7
```

```
Query Match      72.4%; Score 42; DB 4; Length 10;
Best Local Similarity 70.0%; Pred. No. 0.72;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 GKKRRQKK 11
| | | | | | | | | |
Db 1 GKKRRQRR 10

RESULT 6
```

US-09-208-966-16
 Sequence 16, Application US/09208966
 Patent No. 6221355
 GENERAL INFORMATION:
 APPLICANT: Dowdy, Steven F.
 TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
 FILE REFERENCE: 48861/1742
 CURRENT APPLICATION NUMBER: US/09/208,966
 CURRENT FILING DATE: 1998-12-10
 EARLIER APPLICATION NUMBER: 60/082,402
 EARLIER FILING DATE: 1998-04-20
 EARLIER APPLICATION NUMBER: 60/069,012
 EARLIER FILING DATE: 1997-12-10
 NUMBER OF SEQ ID NOS: 57
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 11
 TYPE: PR1
 ORGANISM: human
 US-09-208-966-16

Query Match	72.4%;	Score 42;	DB 4;	Length 11;
Best Local Similarity	70.0%;	Pred. No. 0.78;		
Matches	7;	Conservative	3;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
QY 2 GRKKRQKK 11
    |||||:::
Db 1 GRKKRQRR 10
```

```

RESULT 7
US-09-208-966-52
: Sequence 52, Application US/09208966
: Patent No. 6221355
: GENERAL INFORMATION:
: APPLICANT: Dowdy, Steven F.
: TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
: FILE REFERENCE: 48881/1742
: CURRENT APPLICATION NUMBER: US/09/208,966
: EARLIER FILING DATE: 1998-12-10
: EARLIER FILING DATE: 1998-04-20
: EARLIER APPLICATION NUMBER: 60/082,402
: EARLIER FILING DATE: 1997-12-10
: NUMBER OF SEQ ID NOS: 57
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 52
: LENGTH: 11
: TYPE: PR1
: ORGANISM: human
US-09-208-966-52

```

Query Match	72.48;	Score 42;	DB 4;	Length 11;
Best Local Similarity	70.08;	Pred. No. 0.78;		
Matches	7;	Conservative	3;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy	2	GRKKRRQKK	11
		:::	
Db	2	GRKKRRQRR	11

```

RESULT 8
US-07-694-983-15
; Sequence 15, Application US/07694983
; Patent No. 5432260
;
; GENERAL INFORMATION:
; APPLICANT: Stahl, Philip D.
; TITLE OF INVENTION: HIGH AFFINITY MANNOSE RECEPTOR
; TITLE OF INVENTION: LIGANDS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
;

```

ADDRESS: Irell & Manella
STREET: 545 Middlefield Road, Suite 200
CITY: Menlo Park
STATE: California
COUNTRY: USA

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/694,983

FILING DATE: 19910503
CLASSIFICATION: 530

CLASSIFICATION: C50
ATTORNEY/AGENT INFORMATION:

NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 955
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-327-72
TELEFAX: 415-327-2051

TELEX: 706141

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:

LENGTH: 11 amino acids
TYPE: AMINO ACID

STRANDEDNESS: single
MORPHOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:
NAME/KEY: Peptide

LOCATION: 1
OTHER INFORMATION: /label= Ac-

NAME/REV. D

NAME/REL: REPUBLIC
LOCATION: 11
SUBJECT: INFORMATION

OTHER INFORMATION: /label= -NH2
7-694-983-15

Query Match 63.88; Score 37; DB 1; Length 11;

Best Local Similarity 54.5%; Pred. No. 4.5;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0.

Qy	1	YGRKKRRQKK	11
		: :	
Db	1	YKKKKKKKKKK	11

RESULT 9
 PCT-US95-06077-17
 : Sequence 17, Application PC/TUS9506077
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Immunobiology Research, Institute Inc.
 :
 : TITLE OF INVENTION: Vaccine Interdiction of Extracellular
 : TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus
 : TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar
 : TITLE OF INVENTION: Intercellular Transactivating Strategies
 :
 : NUMBER OF SEQUENCES: 38
 :
 : CORRESPONDENCE ADDRESSES:
 :
 : ADDRESSEE: Howson and Howson
 :
 : STREET: Spring House Corporate Cntr, PO Box 457
 :
 : CITY: Spring House
 :
 : STATE: Pennsylvania
 :
 : COUNTRY: USA
 :
 : ZIP: 19477
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: Floppy disk
 :
 : COMPUTER: IBM PC compatible
 :
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 :
 : SOFTWARE: Patentin Release #1.0, Version #1.25
 :
 : CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06077
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/247,991
FILING DATE: 23-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: IR144PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-06077-17

Query Match 62.1%; Score 36; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RRRRRR 8
|||||
DB 1 RRRRRR 7

RESULT 10
US-08-218-608-3
Sequence 3, Application US/08218608
Patent No. 5607859
GENERAL INFORMATION:
APPLICANT: BIEMANN, KLAUS
APPLICANT: JUHASZ, PETER
TITLE OF INVENTION: METHODS AND PRODUCTS FOR MASS
SPECTROMETRIC MOLECULAR WEIGHT DETERMINATION OF POLYIONIC
TITLE OF INVENTION: ANALYTES EMPLOYING POLYIONIC REAGENTS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,608
FILING DATE: 28-MAR-1994
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: M0656/7013
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: NONE (SYNTHETIC PEPTIDE)
US-08-218-608-3

Query Match 62.1%; Score 36; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 RRRRRR 11
|||||
DB 1 RRRRRR 9

RESULT 11
US-08-357-056-1
Sequence 1, Application US/08357056
Patent No. 5646120
GENERAL INFORMATION:
APPLICANT: SUMNER-SMITH, Martin
APPLICANT: BARNETT, Richard W.
APPLICANT: REID, Lorne S.
APPLICANT: SONENBERG, Nahum
TITLE OF INVENTION: PEPTIDE-BASED INHIBITORS OF HIV
TITLE OF INVENTION: REPLICATION
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/357,056
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/779,735
FILING DATE: 23-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/140 ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-357-056-1

Query Match 62.1%; Score 36; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 RRRRRR 11
|||||
DB 1 RRRRRR 9

RESULT 12

US-08-475-583-1
; Sequence 1, Application US/08475583
; Patent No. 5789531
; GENERAL INFORMATION:
; APPLICANT: SUMNER-SMITH, Martin
; APPLICANT: BARNETT, Richard W.
; APPLICANT: REID, Lorne S.
; APPLICANT: SONENBERG, Nahum
; TITLE OF INVENTION: PEPTIDE-BASED INHIBITORS OF HIV
; TITLE OF INVENTION: REPLICATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,583
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/779,735
; FILING DATE: 23-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/277/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-475-583-1

Query Match 62.1%; Score 36; DB 1; Length 9;
Best local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKRRQKK 11
|||||:::
DB 1 RKKRRQRR 9

RESULT 13
US-08-442-461D-19
; Sequence 19, Application US/08442461D
; Patent No. 5834184
; GENERAL INFORMATION:
; APPLICANT: Harada, Kazuo
; APPLICANT: Martin, Shelley S.
; APPLICANT: Frankel, Alan
; TITLE OF INVENTION: In Vivo Selection of RNA-Binding
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,461D
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 02307U-060500US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ. ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-442-461D-19

Query Match 62.1%; Score 36; DB 2; Length 9;
Best local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKRRQKK 11
|||||:::
DB 1 RKKRRQRR 9

RESULT 14
US-08-893-853-4
; Sequence 4, Application US/08893853
; Patent No. 5891994
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing
; TITLE OF INVENTION: Multiplication of HIV-1
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House
; STATE: PA
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/893,853
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: GGP20USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ. ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-893-853-4

Job time: 118 sec

Query Match 62.1%; Score 36; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKRROKK 11
|||||:
DB 1 RKKRRORR 9

RESULT 15
US-08-564-164A-10
; Sequence 10, Application US/08564164A
; Patent No. 6159947
; GENERAL INFORMATION:
; APPLICANT: Schweighofer, Fablen
; APPLICANT: Tocque, Bruno
; TITLE OF INVENTION: Intracellular Binding Proteins and Use
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,164A
; FILING DATE: 28-DEC-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR94/00714
; FILING DATE: 15-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 93/07241
; FILING DATE: 16-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin F.
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: ST93030-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3816
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-564-164A-10

Query Match 62.1%; Score 36; DB 4; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 RKKRROKK 11
|||||:
DB 1 RKKRRORR 9

Search completed: July 15, 2002, 13:47.49

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:50:01 ; Search time 25.22 Seconds
(without alignments)
110.491 Million cell updates/sec

Title: US-09-712-819A-9

Perfect score: 144

Sequence: 1 YERKKRRQRRRSGSGTDFLTITSSLQAE D 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 5929

Minimum DB seq length: 0

Maximum DB seq length: 29

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	29.2	23	2	S43765
2	39	27.1	29	2	S08555
3	38	26.4	27	1	SRAPAS
4	35	24.3	21	2	PN0082
5	35	24.3	25	1	JC4278
6	35	24.3	25	2	U01617
7	35	24.3	25	2	JC4685
8	35	24.3	25	2	S38425
9	35	24.3	25	2	T06233
10	35	24.3	25	2	T49214
11	34	23.6	27	1	SRAPC
12	33	22.9	15	2	A38304
13	33	22.9	25	2	T38719
14	33	22.2	29	2	S35924
15	31	21.5	21	2	C49042
16	31	21.5	21	2	B49042
17	31	21.5	22	2	I41299
18	31	21.5	25	1	REBY4B
19	31	21.5	28	2	I59477
20	30	20.8	17	2	S26744
21	30	20.1	12	2	S26544
22	29	20.1	20	2	S09720
23	29	20.1	22	2	C64330
24	29	20.1	28	2	I46921
25	29	20.1	28	2	S26234
26	28	19.4	10	2	PH0946
27	28	19.4	16	2	A35552
28	28	19.4	19	2	A33361
29	28	19.4	27	2	S09504

30 28 19.4 28 2 132529
31 27.5 19.1 23 2 S47192
32 27 18.8 12 2 S26546
33 27 18.8 13 2 A53608
34 27 18.8 14 2 PH0915
35 27 18.8 15 2 A49480
36 27 18.8 17 2 T24687
37 27 18.8 18 2 PS0387
38 27 18.8 26 2 T14041
39 26 18.1 10 2 E60787
40 26 18.1 13 1 UNBO
41 26 18.1 15 2 I53284
42 26 18.1 20 2 S06466
43 26 18.1 21 2 PS0146
44 26 18.1 22 2 A28524
45 26 18.1 23 2 I39681

ALIGNMENTS

RESULT 1
S43765
ribosomal protein S19 - evening primrose mitochondrion (fragment)
C:Species: mitochondrion Oenothera villaricae (evening primrose)
C:Date: 10-Dec-1994 #sequence_revision 12-May-1995 #text_change 13-Aug-1999
C:Accession: S43765; S30542
R:Bock, H.; Brennicke, A.; Schuster, W.
Plant Mol. Biol. 24, 811-818, 1994
A:Title: Rps3 and rpl16 genes do not overlap in Oenothera mitochondria: GTG as a pote
A:Reference number: S43765; MUID:94250844
A:Accession: S43765
A:Molecule type: DNA
A:Residues: 1-23 <BOC>
A:Cross-references: EMBL:X69140; NID:913188; PIDN:CAA4893.1; PID:913189
A:Note: the source is designated as Oenothera berteriana
C:Genetics:
A:Gene: rps19
A:Genome: mitochondrion
C:Superfamily: Escherichia coli ribosomal protein S19
C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 29.2% Score 42; DB 2; Length 23;
Best Local Similarity 53.3% Pred. No. 30;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Oy 1 YERKKRRQRRRSGSG 15
Db 3 FTKKRRSRKTWIGSG 17

RESULT 2
S08555
ribosomal protein L15 [validated] - Halobacterium salinarum (fragment)
N:Alternate names: ribosomal protein HL16
C:Species: Halobacterium salinarum
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jul-2000
C:Accession: S08555
R:Matheson, A.T.; Yaguchi, M.; Christensen, P.; Rollin, C.F.; Hasnain, S.
Can. J. Biochem. Cell Biol. 62, 426-433, 1984
A:Title: Purification, properties, and N-terminal amino acid sequence of certain 50S
A:Reference number: S07437; MUID:84282108
A:Accession: S08555
A:Molecule type: protein
A:Residues: 1-29 <MAT>
A:Note: the source is designated as Halobacterium cutirubrum
A:Note: the protein is designated as ribosomal protein L16
C:Superfamily: rat ribosomal protein L27a
C:Keywords: protein biosynthesis; ribosome

Query Match 27.1% Score 39; DB 2; Length 29;

Best Local Similarity 66.7%; Pred. No. 93;
Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
OY 4 KRRROR--RRSGSGT 16
||||| |
Db 3 KRRRORSGRTHGGGT 17

RESULT 3
SRAPAS
protamine A - stellate sturgeon
N:Alternate names: stellin
C:Species: Acipenser stellatus (stellate sturgeon)
C>Date: 28-Feb-1980 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997
C:Accession: A02666
R:Yulikova, E.P.; Rybkin, V.K.; Silaev, A.B.
Bioorg. Khim. 5, 5-10, 1979
A:Title: The primary structure of stellin A.
A:Reference number: A02666
A:Accession: A02666
A:Molecule type: protein
A:Residues: 1-27 <YUL>
A:Note: article in Russian with English abstract
C:Superfamily: protamine Y2
C:Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match 26.4%; Score 38; DB 1; Length 27;
Best Local Similarity 70.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 KRRRORRRSG 13
|:|:|:| |
Db 13 KRRRRRRRHHG 22

RESULT 4
PN0082
sperm chromatin protein 12-2 - Argentinian shortfin squid (fragment)
C:Species: Illex argentinus (Argentinian shortfin squid)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PN0082
R:Osadchuk, L.A.; Levina, N.B.; Telezhinskaya, I.N.; Khrapunov, S.N.; Berdyshev, G.D.; A
Bioorg. Khim. 16, 448-455, 1990
A:Title: Primary structure of main nuclear protein from headleg mollusk *Illex argentinus*
A:Reference number: PN0081; MUID:90329035
A:Accession: PN0082
A:Molecule type: protein
A:Residues: 1-21 <OSA>
A:Experimental source: sperm
C:Superfamily: sperm histone
C:Keywords: sperm

Query Match 24.3%; Score 35; DB 2; Length 21;
Best Local Similarity 58.3%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRRRRRRRRSGS 14
|:|:|:| |
Db 1 RRRRRRRRRRRS 12

RESULT 5
JC4278
ribosomal protein L41, cytosolic (similarity) - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 01-Sep-2000
C:Accession: JC4278
R:Chan, Y.L.; Olivera, J.; Wool, I.G.
Biochem. Biophys. Res. Commun. 214, 810-818, 1995
A:Title: The primary structures of rat ribosomal proteins L4 and L41.
A:Reference number: JC4277; MUID:96024571

A:Accession: JC4278
A:Molecule type: mRNA
A:Residues: 1-25 <CHA>
A:Cross-references: EMBL:X82550; NID:9575583; PIDN:CAAT899.1; PID:9575384
C:Superfamily: rat ribosomal protein L41
C:Keywords: ribosome

Query Match 24.3%; Score 35; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
|:|:|:| |
Db 14 KRRKKRRORS 24

RESULT 6
J01617
ribosomal protein L41e - human
N:Alternate names: HGI2 protein; ribosomal protein YL41
C:Species: Homo sapiens (man)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: J01617; JC5659; S31691
R:Klaudiny, J.; von der Kammer, H.; Scheit, K.H.
Biochem. Biophys. Res. Commun. 187, 901-906, 1992
A:Title: Characterization by cDNA cloning of the mRNA of a highly basic human protein
A:Reference number: J01617; MUID:92412140
A:Accession: J01617
A:Molecule type: mRNA
A:Residues: 1-25 <KLA>
A:Cross-references: EMBL:Z12962; NID:936135; PIDN:CAAT8306.1; PID:936136
R:Lee, J.H.; Kim, J.M.; Kim, M.S.; Lee, Y.T.; Marshak, D.R.; Bae, Y.S.
Biochem. Biophys. Res. Commun. 238, 462-467, 1997
A:Title: The highly basic ribosomal protein L41 interacts with the beta subunit of pr
A:Reference number: JC5659; MUID:97446005
A:Accession: JC5659
A:Molecule type: mRNA
A:Residues: 1-25 <LEB>
A:Cross-references: EMBL:Z12962; NID:936135; PIDN:CAAT8306.1; PID:936136
C:Comment: This protein stimulates phosphorylation of the beta chain of DNA topoisome
C:Superfamily: rat ribosomal protein L41

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
|:|:|:| |
Db 14 KRRKKRRORS 24

RESULT 7
JC4685
ribosomal protein L41 - cat
C:Species: Felis silvestris catus (domestic cat)
C>Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 18-Aug-2000
C:Accession: JC4685
R:Starkey, C.R.; Menon, R.P.; Prabhu, S.; Levy, L.S.
Biochem. Biophys. Res. Commun. 220, 648-652, 1996
A:Title: Primary sequence and evolutionary conservation of ribosomal protein genes fr
A:Reference number: JC4685; MUID:96183078
A:Accession: JC4685
A:Molecule type: mRNA
A:Residues: 1-25 <SRA>
A:Cross-references: GB:U22229; NID:9950108; PIDN:AA01667.1; PID:9950109
A:Experimental source: thymic lymphoma
C:Genetics:
A:Gene: rpl41
C:Superfamily: rat ribosomal protein L41

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRQRRS 12
:|:|:|:|:|
Db 14 KKKRRKMRORS 24

RESULT 8

S38425 ribosomal protein GL41 - upland cotton

C:Species: Gossypium hirsutum (upland cotton)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Aug-2000

C:Accession: S38425

R:Turley, R.B.; Ferguson, D.L.; Meredith, W.R.
submitted to the EMBL Data Library, October 1993

A:Reference number: S38425

A:Accession: S38425

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-25 <TUR>

A:Cross-references: EMBL:X75423; NID:g407800; PIDN:CA53175.1; PID:g407801

C:Superfamily: rat ribosomal protein L41

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRQRRS 12
:|:|:|:|:|
Db 14 KKKRRKMRORS 24

RESULT 9

T06233 ribosomal protein L41 - soybean (fragment)

C:Species: Glycine max (soybean)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 18-Aug-2000

C:Accession: T06233

R:Mahalingam, R.; Knap, H.T.
submitted to the EMBL Data Library, February 1998

A:Reference number: Z15555

A:Accession: T06233

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-25 <MAH>

A:Cross-references: EMBL:AF047051; NID:g2905777; PIDN:AAC0557.1; PID:g2905778

A:Experimental source: cultivar Essex; root

C:Superfamily: rat ribosomal protein L41

C:Keywords: protein biosynthesis; ribosome

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRQRRS 12
:|:|:|:|:|
Db 14 KKKRRKMRORS 24

RESULT 10

T49214

ribosomal protein GL41-like - Arabidopsis thaliana

N:Alternate names: protein F27K19.200

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000

C:Accession: T49214

R:Benes, V.; Wurmbech, E.; Drzonek, H.; Ansoorge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
submitted to the Protein Sequence Database, April 2000

A:Reference number: Z25014

A:Accession: T49214
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-25 <BEN>

A:Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.200

A:Experimental source: cultivar Columbia; BAC clone F27K19

C:Genetics:

A:Gene: ATSP:F27K19.200

A:Map position: 3

A:Introns: 4/3

C:Superfamily: rat ribosomal protein L41

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRQRRS 12
:|:|:|:|:|
Db 14 KKKRRKMRORS 24

RESULT 11

SRAPC

protamine B - Russian sturgeon

N:Alternate names: sturine

C:Species: Acipenser gueldenstaedti (Russian sturgeon)

C>Date: 12-Aug-1981 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997

C:Accession: A02665

R:Yulikova, E.P.; Evseenko, L.K.; Baratova, L.A.; Belyanova, L.P.; Rybin, V.K.; Silae

Bloorg, Khm. 2, 1613-1617, 1976

A:Reference number: A02665

A:Accession: A02665

A:Molecule type: protein

A:Residues: 1-27 <YUL>

A:Superfamily: protamine Y2

C:Keywords: Chromosomal protein; DNA binding; spermatogenesis

Query Match 23.6%; Score 34; DB 1; Length 27;
Best Local Similarity 60.0%; Pred. No. 4e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KKRQRRRSG 13
:|:|:|:|:|
Db 11 QRRRRRRHG 20

RESULT 12

A38304

heterogeneous ribonuclear particle protein A1 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Oct-1997

C:Accession: A38304

R:Kumar, A.; Casas-Finet, J.R.; Luneau, C.J.; Karpel, R.L.; Merrill, B.M.; Williams,

J. Biol. Chem. 265, 17094-17100, 1990

A:Title: Mammalian heterogeneous nuclear ribonucleoprotein A1. Nucleic acid binding p

A:Reference number: A38304; MUID:91009136

A:Accession: A38304

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-15 <KUM>

C:Superfamily: helix-stabilizing protein; ribonucleoprotein repeat homology

Query Match 22.9%; Score 33; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 8 QRRSGSG 15
:|:|:|:|:|
Db 8 QRRSGSG 15

RESULT 13

T38719

ribosomal protein L41 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000

C:Accession: T38719; T38731; T39176

R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z21807

A:Accession: T38719

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-25 <MUR>

A:Cross-references: EMBL:Z69369; PIDN:CAB40187.1; GSPDB:GN00066

A:Experimental source: strain 972h-; cosmid C3F10

R:Gentile, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z21797

A:Accession: T38731

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-25 <GEN>

A:Cross-references: EMBL:Z99167; PIDN:CAB40152.1; GSPDB:GN00066; SPDB:SPAC3G6.13c

A:Experimental source: strain 972h-; cosmid C3G6

R:Bothe, G.; Pohl, T.; McDougall, R.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21833

A:Accession: T39176

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-25 <BOT>

A:Cross-references: EMBL:AL109738; PIDN:CAB52162.1; GSPDB:GN00066; SPDB:SPAC8F11.01c

A:Experimental source: strain 972h-; cosmid c8F11

A:Gene: SPAC31F10.18c; SPAC3G6.13c; SPAC3F10.18c; SPAC8F11.01c

A:Map position: 1

A:introns: 24/2

C:Superfamily: rat ribosomal protein L41

Query Match 22.9%; Score 33; DB 2; Length 25;
 Best Local Similarity 54.5%; Pred. No. 5.1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ERKRRRRRS 12

DB 14 KRRRRRRARS 24

RESULT 14

S35924

T-cell receptor gamma chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 30-Jun-2001

C:Accession: S35924

R:Mathioudakis, G.; Platoucas, C.D.

submitted to the EMBL Data Library, May 1993

A:Description: An alternative splicing between V-gamma, J-gamma2.3 and C-gamma2 gene seg

A:Reference number: S32764

A:Accession: S35924

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-29 <MAT>

A:Cross-references: EMBL:Z22684

C:Keywords: T-cell receptor

DB 3 KGMKRLFGSGTTLVVT 18
 | : | | | : |

RESULT 15

C49042

Ig heavy chain V region, anti-DNA monoclonal G2a antibody - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 16-Aug-1996

C:Accession: C49042

R:Takl, S.; Hirose, S.; Kinoshita, K.; Nishimura, H.; Shimamura, T.; Hamuro, J.; Shlr

Eur. J. Immunol. 22, 987-992, 1992

A:Title: Somatic mutation of IgG anti-DNA antibody clonally related to germ-line enco

A:Reference number: A49042; MUID:92201320

A:Contents: NZB X NZW F1, hybridoma BW9-45

A:Accession: C49042

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-21 <TAK>

A:Note: sequence extracted from NCBI backbone (NCBIP:92267)

C:Keywords: heterotetramer; immunoglobulin

Query Match 21.5%; Score 31; DB 2; Length 21;
 Best Local Similarity 66.7%; Pred. No. 8e+02;
 Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

OY 13 GSGTDFTLTSS 24

DB 12 GGGT--TLTVSS 21

Search completed: July 15, 2002, 13:52:41

Job time: 160 sec

Query Match 22.2%; Score 32; DB 2; Length 29;
 Best Local Similarity 43.8%; Pred. No. 8e+02;
 Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 6 RRORRRSGGTFTLT 21

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:51:47 ; Search time 12.84 Seconds

(without alignments)
87,451 Million cell updates/sec

Title: US-09-712-819a-9

Perfect score: 144

Sequence: 1 YERKRRRRSGSGTFTLTISLQAEQ 29

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1766

Minimum DB seq length: 0

Maximum DB seq length: 29

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	27.1	29	1	RL15_HALCU
2	38	26.4	27	1	PRTA_ACTIST
3	35	24.3	25	1	RL41_ARATH
4	35	24.3	25	1	RL41_HUMAN
5	35	24.3	25	1	RL41_SOYBN
6	34	23.6	27	1	PR15_ACTIBN
7	33	22.9	25	1	RL41_SCHPO
8	32	22.2	25	1	RL41_AGABI
9	32	22.2	25	1	RL41_QUEBU
10	31	21.5	25	1	RL41_YEAST
11	29	20.1	22	1	RL41_METUA
12	28	19.4	27	1	HM12_CABEL
13	28	19.4	29	1	TL16_SPIOI
14	27.5	19.1	29	1	GLUC_LAMFL
15	27	18.8	13	1	NEUT_CAVPO
16	26	18.1	20	1	PSBW_WHEAT
17	25	17.4	22	1	RL18_HALME
18	25	17.4	23	1	RL18_HALVO
19	25	17.4	24	1	HEMO_LINRE
20	25	17.4	24	1	SCRK_LACIA
21	25	17.4	26	1	RL18_HALHA
22	24	16.7	29	1	GLUC_DIDMA
23	24	16.7	15	1	GLUC_RABIT
24	23	16.0	15	1	UC06_MAIZE
25	23	16.0	24	1	VGJ_BPAL3
26	23	16.0	28	1	RS11_THETH
27	23	16.0	25	1	VG9_SPV4
28	23	16.0	29	1	GLUC_CALMI
29	23	15.3	11	1	CA41_LITCI
30	22	15.3	12	1	UR2A_CATCO
31	22	15.3	20	1	TENB_ACTTE
32	22	15.3	23	1	POQA_KLEPN
33	22	15.3	29	1	RP54_CLOKL

34	21	14.6	9	1	DL_NERNO	P24816 nephrops no
35	21	14.6	9	1	FAR6_CALVO	P41861 calliphora
36	21	14.6	11	1	CA42_LITCI	P82092 litoria cit
37	21	14.6	16	1	RIPK_TRIKI	P16093 trichosanth
38	21	14.6	18	1	YAA5_RHOPA	002005 rhodopsendo
39	21	14.6	19	1	LANA_ACTIG	P56650 actinoplane
40	21	14.6	20	1	TENA_ACTTE	P30833 actinia ten
41	21	14.6	22	1	TF2G_PIG	P20461 sus scrofa
42	21	14.6	22	1	MOT1_CANFA	P19863 canis fam1
43	21	14.6	22	1	TL11_SPIOI	P82657 spinacia ol
44	21	14.6	25	1	ANDT_ANDAU	P56684 androctonus
45	21	14.6	25	1	VGJ_BPC4	P03652 bacterioph

ALIGNMENTS

RESULT 1						
ID	RL15_HALCU	STANDARD;	PRT;	29	AA.	
AC	P05971;					
DT	01-NOV-1988 (Rel. 09, Created)					
DT	01-NOV-1988 (Rel. 09, Last sequence update)					
DT	30-MAY-2000 (Rel. 39, Last annotation update)					
DE	50S ribosomal protein L15P (HL16) (Fragment).					
GN	RP.L15P.					
OS	Halobacterium cutludrum.					
OC	Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;					
OC	Halobacterium.					
OX	NCBI_TaxID=2242;					
RN	[1]					
RP	SEQUENCE.					
RA	MEDLINE=84282108; PubMed=6467081;					
RX	Matheson A.T., Yaguchi M., Christensen P., Rollin C.F., Hasnain S.;					
RT	*Purification, properties, and N-terminal amino acid sequence of					
RT	certain 50S ribosomal subunit proteins from the archaebacterium					
RT	Halobacterium cutludrum.*					
RL	Can. J. Biochem. Cell Biol. 62:426-433(1984).					
CC	-1- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.					
DR	PIR: S08555; S08555.					
DR	InterPro: IPR001196; Ribosomal_L15.					
DR	ProSITE: PS00475; RIBOSOMAL_L15; PARTIAL.					
KW	Ribosomal protein.					
FT	NON_TER					
SQ	SEQUENCE 29 AA; 3170 MW; 73D13D831B4BE790 CRC64;					
QY	4 KRRRR-RRSGGT 16					
Db	3 KRRRRSGRTGGGT 17					
RESULT 2						
ID	PRTA_ACTIST	STANDARD;	PRT;	27	AA.	
AC	P02324;					
DT	21-JUL-1986 (Rel. 01, Created)					
DT	21-JUL-1986 (Rel. 01, Last sequence update)					
DT	01-NOV-1995 (Rel. 32, Last annotation update)					
DE	Protamines A and C (Stellin A and C).					
OC	Acipenser stellatus (Sevruga).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;					
OC	Acipenserinae; Acipenser.					
OX	NCBI_TaxID=7903;					
RN	[1]					
RP	SEQUENCE (STELLIN A).					
RA	Yulikova E.P., Rybin V.K., Silaev A.B.;					
RT	*The primary structure of stellin A.*;					

RL Bloorg. Khim. 5:5-10(1979).
 RN [2]
 RP SPECIES-A.thaliana; STRAIN-CV. COLUMBIA;
 RX MEDLINE-94019256; PubMed-1339950;
 RA Rybin V.K., Revina L.P., Baratova L.A.;
 RT "Comparison of amino acid sequences of sturgeon triptamines using
 CC proteinases from *Acipenser stellatus* gonads as an example.";
 CC Mol. Biol. (Mosk) 26:300-306(1992).
 CC -1- FUNCTION: PROTEINASES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: Testis.
 DR PIR: A02656; SRAPS.
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
 KW Testis; DNA condensation; Nuclear protein.
 FT VARIANT 1 1 MISSING (IN STELLIN C).
 SO SEQUENCE 27 AA; 3532 MW; 3EAA5B950FA8658 CRC64;
 Query Match 26.4%; Score 38; DB 1; Length 27;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Oy 4 KRRRRRRRG 13
 Db 13 KRRRRRRRG 22
 RESULT 3
 RL41_ARATH STANDARD; PRT; 25 AA.
 ID RL41_ARATH
 AC P35015;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L41
 GN (RPL41-A OR AT1G56045 OR T6H22.15) AND (RPL41-B OR AT3G08520 OR
 GN T6G24.5 OR T6G24.5) AND (RPL41-C OR AT3G11120 OR F9F8.7) AND
 GN (RPL41-D OR AT3G56020 OR F27K19-200).
 OS Arabidopsis thaliana (Mouse-ear cress),
 OS Nicotiana tabacum (Common tobacco),
 OS Gossypium hirsutum (Upland cotton),
 OS Pisum sativum (Garden pea),
 OS Hordeum vulgare (Barley), and
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702, 4097, 3635, 3888, 4513, 4530;
 RN [1]
 RN SEQUENCE FROM N.A. (AT1G56045).
 RP SPECIES-A.thaliana; STRAIN-CV. COLUMBIA;
 RX MEDLINE-21016719; PubMed-11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malt R., Marshall A.,
 RA Miltischer J., Miranda M., Nguyen M., Nieman W.C., Osborne B.I.,
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
 RA Sun H., Tallon L.J., Tanbarga G., Toriumi M.J., Town C.D.,
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
 RL thaliana."; Nature 408:816-820(2000).

RN [2]
 RP SPECIES-A.thaliana; STRAIN-CV. COLUMBIA;
 RX MEDLINE-21016720; PubMed-11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansgore W., Unseld M.,
 RA Farmanan B., Valle G., Bloeker H., Perez-Alonso M., Obermayer B.,
 RA Deiseny M., Bouty M., Grivell L.A., Maché R., Pulgomech P.,
 RA De Simone V., Choisme N., Artiguenave F., Robert C., Brotlier P.,
 RA Winkler P., Cattolico L., Weissenbach J., Savin W., Quetier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Berns V.,
 RA Wurmbech E., Drzonek H., Erle H., Jordan R., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Bradt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Topo S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Scharfe M., Schoen O., Barues M., Terol J., Climent J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemin D.,
 RA Cooke R., Landie M., Berger-Liauro C., Purrelli B., Masny D.,
 RA de Haan M., Maare A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argilou A., Flores M., Liguori R., Vitale D.,
 RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Malt R., Uterback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Malt R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
 RA Preuss D., Lin X., Nieman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
 RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakayama S., Nakazaki N., Shinozaki S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 RL thaliana."; Nature 408:822(2000).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP SPECIES-A.thaliana; STRAIN-CV. COLUMBIA;
 RC Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Gernlind P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamliya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin X., Liu S.X., Miranda M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakuraki T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (Sep-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RP SPECIES-N.tabacum; STRAIN-CV. SRL; TISSUE-Leaf;
 RA Zhou X.R.;
 RL Submitted (Apr-1995) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A.
 RP SPECIES-G.hirsutum; STRAIN-CV. DELTAPINE 62;
 RX MEDLINE-95062728; PubMed-7972506;
 RA Tuttle R.B., Ferguson D.L., Meredith W.R.;
 RT "Isolation and characterization of a cDNA encoding ribosomal protein
 RL L41 from cotton (*Gossypium hirsutum* L.).";
 RL Plant Physiol. 105:1449-1450(1994).
 RN [6]
 RN SEQUENCE FROM N.A.
 RP SPECIES-P.sativum; STRAIN-CV. LITTLE MARVEL; TISSUE-Root tip;
 RC Woo H.H.;
 RL Submitted (Oct-1995) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RN SEQUENCE FROM N.A.
 RP SPECIES-H.vulgare; STRAIN-CV. Boml; TISSUE-Endosperm;
 RC Raasmussen S.R.;
 RT "Barley L41 ribosomal protein from immature endosperm.";
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 RN [8]
 RN SEQUENCE FROM N.A.
 RP SPECIES-O.sativa; STRAIN-CV. Nipponbare;

RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
 RA Zisman V., Pal G., Bowman C.L., Fujii C.Y., Vancken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Frazer C.M.,
 RT Oryza sativa chromosome 3 BAC OSJNB009119 genomic sequence."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE LA1E FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AC009894; AF02845.1; -
 DR EMBL: AC074395; AAG50829.1; -
 DR EMBL: AC009991; AAF01511.1; -
 DR EMBL: AL163832; CAB87856.1; -
 DR EMBL: AY058053; AAL24161.1; -
 DR EMBL: U26255; AAG7297.1; -
 DR EMBL: X75423; CAA53175.1; -
 DR EMBL: L47967; AAA79268.1; -
 DR EMBL: AJ001160; CAA04364.1; -
 DR EMBL: AC084320; AAK09215.1; -
 DR PIR: S38425; S38425.
 KW Ribosomal protein.
 SQ SEQUENCE 25 AA; 3428 MW; 9AD629D4293C039E CRC64;
 OY 2 ERKKRRORRRS 12
 DB 14 KKKRKKMRRS 24
 ID RL41_HUMAN STANDARD: PRT: 25 AA.
 AC P28731.
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L41 (HG12).
 CN RP41.
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat).
 OS Felis silvestris catus (Cat).
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606, 10090, 10116, 9685, 31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human: TISSUE-Ovary;
 RX MEDLINE=92412140; PubMed=1326959;
 RA Klaidiny J., von der Kammer H., Scheit K.H.:
 RT "Characterization by cDNA cloning of the mRNA of a highly basic human
 RT protein homologous to the yeast ribosomal protein YL1.";
 RL Biochem. Biophys. Res. Commun. 187:901-906(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human:
 RX MEDLINE=97446005; PubMed=9299532;
 RA Lee J.-H., Kim J.-M., Kim M.-S., Lee Y.-T., Marshak D.R., Bae Y.-S.:
 RT "The highly basic ribosomal protein L41 interacts with the beta
 RT subunit of protein kinase CKII and stimulates phosphorylation of DNA
 RT topoisomerase IIalpha by CKII.";

RL Biochem. Biophys. Res. Commun. 238:462-467(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human:
 RA Co H., Miyado K., Hasewa H., Taniguchi S.:
 RT "Characterization of human ribosomal protein L41 genomic structure."
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human:
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levasilho M.H., Leverisha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prichaltingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durlin R.M., Bentley D.R., Beck S.,
 RA Rogers J.:
 RT "The DNA sequence and comparative analysis of human chromosome 20."
 RL Nature 414:865-871(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rat: STRAIN-SPRAGUE-DAWLEY; TISSUE-Liver;
 RA Chan Y.-L., Wool I.G.:
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Mouse: STRAIN-C57BL/6;
 RA Koeha D., Anderson E., Botcherby M., Jordan B.,
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES-F.s. catus;
 RX MEDLINE=96183078; PubMed=8607819;
 RA Starkey C.R., Menon R.P., Pradhu S., Levy L.S.:
 RT "Primary sequence and evolutionary conservation of ribosomal protein
 RT genes from the domestic cat."
 RL Biochem. Biophys. Res. Commun. 220:648-652(1996).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC SPECIES-F. rubripes;
 RX MEDLINE=99177347; PubMed=10077531;
 RA Gellner K., Brenner S.:
 RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
 RT rubripes."
 RL Genome Res. 9:251-258(1999).
 CC -1- FUNCTION: INTERACTS WITH THE BETA SUBUNIT OF PROTEIN KINASE CKII
 CC AND STIMULATES PHOSPHORYLATION OF DNA TOPOISOMERASE II ALPHA BY
 CC CKII.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

```

CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: Z12962; CAA78306.1; -
DR  EMBL: AF026844; ABB82715.1; -
DR  EMBL: AB010874; BAA31508.1; -
DR  EMBL: AL035562; CAB46594.1; -
DR  EMBL: X82550; CAA57899.1; -
DR  EMBL: U93862; AAB52254.1; -
DR  EMBL: U22229; AAB01667.1; -
DR  EMBL: AF056116; AAC34393.1; -
DR  PIR: J01617; J01617.
DR  PIR: S31691; S31691.
KW  Ribosomal protein.
SQ  SEQUENCE 25 AA: 3456 MW; 9AD629C69ECC039E CRC64;

Query Match
Best Local Similarity 24.3%; Score 35; DB 1; Length 25;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  2 ERKKRRRRRS 12
Db  14 KRRRRRRRS 24

RESULT  5
ID  RL41_SOYBN STANDARD; PRT; 25 AA.
AC  049224;
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DE  60S ribosomal protein L41.
GN  RPL41.
OS  Glycine max (Soybean).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eustosios 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX  NCBI_TaxID=3847;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CV. ESSEX; TISSUE=Root;
RA  Woo H.-H.;
RL  Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: BELONGS TO THE LA1E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF047051; AAC03557.1; -
DR  EMBL: AF047051; AAC03557.1; -
KW  Ribosomal protein.
SQ  SEQUENCE 25 AA: 3130 MW; AD933FB8E49EDD6E CRC64;

Query Match
Best Local Similarity 24.3%; Score 35; DB 1; Length 25;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY  2 ERKKRRRRRS 12
Db  14 KRRRRRRRS 24

RESULT  6
ID  PRTB_ACIGU STANDARD; PRT; 27 AA.
AC  P02323;

```

```

DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-DEC-1992 (Rel. 24, Last annotation update)
DE  Protamine B (Sturine B) (Stellin B).
OS  Acipenser guineensis (Caspian sturgeon) (Russian sturgeon), and
05  Acipenser stellatus (Sevruga).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OX  NCBI_TaxID=7902, 7903;
RN  [1]
RP  SEQUENCE.
RC  SPECIES-A.guineensis;
RA  Yulikova E.P., Eysenko L.K., Baratova L.A., Belyanova L.P.,
RA  Rybin V.K., Siliev A.B.;
RT  "The primary structure of sturine B, a protamine from Caspian
RT  sturgeon."
RL  Bioorg. Khim. 2:1613-1617(1976).
RN  [2]
RP  SEQUENCE.
RC  SPECIES-A.stellatus;
RA  Rybin V.K., Yulikova E.P.;
RL  Khim. Prirod. Soedin. 5:700-704(1979).
CC  -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC  SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPLEX
CC  SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC  -1- SUBCELLULAR LOCATION: Nuclear.
CC  -1- TISSUE SPECIFICITY: TESTIS.
DR  PIR: A02665; SRAPC.
KW  Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW  Testis; DNA condensation; Nuclear protein.
SQ  SEQUENCE 27 AA: 3707 MW; E300B46B1737EC80 CRC64;

Query Match
Best Local Similarity 23.6%; Score 34; DB 1; Length 27;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY  4 KRRRRRRRS 13
Db  11 QRRRRRRRS 20

RESULT  7
ID  RL41_SCHPO STANDARD; PRT; 25 AA.
AC  09Y710;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  60S ribosomal protein L41.
GN  (RPL41A OR SPAC3G6.13c) AND (RPL41B OR SPAC3JF10.18c OR SPAC8F11.01c).
OS  Schizosaccharomyces pombe (fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OC  Schizosaccharomyces.
OX  NCBI_TaxID=4896;
RN  [1]
RP  SEQUENCE FROM N.A. (RPL41A).
RC  STRAIN-972;
RA  Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL  Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A. (RPL41B).
RC  STRAIN-972;
RA  Bothe G., Pohl T., McDougall R., Rajandream M.A., Barrell B.G.;
RL  Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC  -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L41 IN S.POMBE.
CC  -1- SIMILARITY: BELONGS TO THE LA1E FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 299167; CAB40152.1; -
DR EMBL; 269369; CAB40187.1; -
DR EMBL; AL109738; CAB52162.1; -
KW Ribosomal protein; Multigene family.
SQ SEQUENCE 25 AA; 3411 MW; 9EC629DD9DECC0535 CRC64;

Query Match
Best Local Similarity 22.2%; Score 32; DB 1; Length 25;
Pred. No. 78;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
Db 14 KRRRRRRARS 24

RESULT 8
RL41_AGABI STANDARD; PRT; 25 AA.
AC P78569;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L41.
GN RL41.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OC NCBI_TaxID=5341;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HORST U1;
RX MEDLINE=97111990; PubMed=8953726;
RA Sonnenberg A.S.M., de Groot P.W.J., Schaap P.J., Baars J.J.P.,
RA Visser J., van Griensven L.J.L.D.;
RT "Isolation of expressed sequence tags of Agaricus bisporus and their
RT assignment to chromosomes."
RL Appl. Environ. Microbiol. 62:4542-4547(1996).
CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X94764; CAA64390.1; -
KW Ribosomal protein.
SQ SEQUENCE 25 AA; 3413 MW; A97629D0143C04EA CRC64;

Query Match
Best Local Similarity 22.2%; Score 32; DB 1; Length 25;
Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 KRRRRRRRS 12
Db 15 KRRRRRRARS 24

RESULT 9
RL41_QUESTU STANDARD; PRT; 25 AA.
ID RL41_QUESTU
AC O82713;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

```

DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L41.
GN RPL41.
OS Quercus suber (Cork oak).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eurosid II; Fagales; Fagaceae; Quercus.
OC NCBI_TaxID=58331;
RN [1]
RP SEQUENCE FROM N.A.
RA Huguet G., Pla M., Verdaguer D., Molinas M.;
RT "Ribosomal proteins in Quercus suber."
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001347; CAA04691.1; -
KW Ribosomal protein.
SQ SEQUENCE 25 AA; 3337 MW; 6BC629DD9EDB3383 CRC64;

Query Match
Best Local Similarity 22.2%; Score 32; DB 1; Length 25;
Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
Db 14 KRRRRRRARS 24

RESULT 10
RL41_YEAST STANDARD; PRT; 25 AA.
ID RL41_YEAST
AC P05746;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L41 (YL41) (YL47).
GN (RPL41A OR RPL47A OR YL41A OR YDL184C OR D1290) AND
GN (RPL41B OR RPL47B OR YL41B OR YDL133BC).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. (RPL41A AND RPL41B).
RX MEDLINE=90254826; PubMed=2187623;
RA Suzuki K., Hashimoto T., Otake E.;
RT "Yeast ribosomal proteins: XI. Molecular analysis of two genes
RT encoding YL41, an extremely small and basic ribosomal protein, from
RT Saccharomyces cerevisiae."
RL Curr. Genet. 17:185-190(1990).
RN [2]
RP SEQUENCE FROM N.A. (RPL41A).
RC STRAIN=5288C / FY1679.
RX MEDLINE=96021607; PubMed=8533471;
RA Verhasselt P., Voet M., Volckaert G.;
RT "New open reading frames, one of which is similar to the nifv gene of
RT Azotobacter vinelandii, found on a 12.5 kbp fragment of chromosome IV
RT of Saccharomyces cerevisiae."
RL Yeast 11:961-966(1995).
RN [3]
RP SEQUENCE FROM N.A. (RPL41B).
RC STRAIN=5288C / FY1679.
RX MEDLINE=97127826; PubMed=8972577;
RA Woelfl S., Haneman V., Saluz H.P.;

```

```

RT *Analysis of a 26,756 bp segment from the left arm of yeast
RT chromosome IV.*
RL Yeast 12:1549-1554(1996).
RN [4]
RP SEQUENCE.
RA Ohtaka E., Higo K.-I., Itoh T.
RT *Yeast ribosomal proteins: VIII. Isolation of two proteins and
RT sequence characterization of twenty-four proteins from cytoplasmic
RT ribosomes.*
RL Mol. Gen. Genet. 195:544-546(1984).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L41 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X16065; CAA34201.1; -
DR EMBL: X16066; CAA34202.1; -
DR EMBL: X83276; CAA58262.1; -
DR EMBL: 274232; CAA98759.1; -
DR EMBL: 274181; CAA98705.1; -
DR EMBL: 274182; CAA98706.1; -
DR EMBL: X96876; CAA65626.1; -
DR PIR: S22247; R6BY4A.
DR PIR: S22246; R6BY4B.
DR SGD: S0002343; RPL41A.
DR SGD: S0002293; RPL41B.
DR Ribosomal protein; Multigene family.
KW SEQUENCE 25 AA; 3337 MW; BD2629DD9ED85381 CRC64;
SQ

```

Query Match 21.5%; Score 31; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

OY 2 ERKRRRRRRS 12
DB 14 KRRRRKRRARS 24

```

RESULT 11
RL41_METJA STANDARD; PRT; 22 AA.
AC P54025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L41E.
GN RPL41E OR M70242.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
CC Methanococcus
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kervatage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.*
RL Science 273:1058-1073(1996).

```

CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U67480; AAB98230.1; -
DR TIGR: MJ0242; -
KW Ribosomal protein; Complete proteome.
KW SEQUENCE 22 AA; 3089 MW; F948D74D4EC98CB CRC64;
SQ

```

Query Match 20.1%; Score 29; DB 1; Length 22;
Best Local Similarity 41.7%; Pred. No. 2.6e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```

OY 3 RKRRRRRRSGS 14
DB 11 RLKRRKRRARS 22

```

RESULT 12
HM12_CAEEL STANDARD; PRT; 27 AA.
AC P17487;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Homeobox protein ceh-12 (Fragment).
GN Ceh-12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90245646; PubMed=1970877;
RA Schaller D., Wiltmann C., Splicher A., Mueller F., Tobler H.;
RT "Cloning and analysis of three new homeobox genes from the nematode
RT Caenorhabditis elegans.";
RL Nucleic Acids Res. 18:2033-2036(1990).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X17076; CAA34928.1; -
DR PIR: S09504; S09504.
DR HSSP: P02836; IENH.
DR TRANSFAC: T02987; -
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR PROSITE: PS00027; HOMEBOX_1; PARTIAL.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA_BIND 1 16 HOMEBOX.
FT NON_TER 27
FT NON_TER 27
SQ SEQUENCE 27 AA; 3182 MW; FDAF5D16F3870A5 CRC64;

Query Match 19.4%; Score 28; DB 1; Length 27;
Best Local Similarity 31.8%; Pred. No. 4.6e+02;
Matches 7; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 3 RKKRRRRSGSGTDTLTSS 24
 : : : : : : : : : : : : : : : :
 Db 6 QNRMRNKRCPSSPTIGSTSS 27

RESULT 13
 T1L6_SPIOL
 ID T1L6_SPIOL STANDARD; PRT; 29 AA.
 AC P81834;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thylakoid lumenal 16.5 kDa protein (P16.5) (Fragment).
 OS Spinacia oleracea (Spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
 OX NCBI_TaxID=3562;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Leaf;
 RX MEDLINE=98175931; PubMed=9506969;
 RA Kieselbach T., Hagman A., Andersson B., Schroeder W.P.;
 RT "The thylakoid lumen of chloroplasts - isolation and
 RT characterization";
 RL J. Biol. Chem. 273:6710-6716(1998).
 KW -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
 CC Chloroplast; Thylakoid.
 FT NON_TER 29
 SQ SEQUENCE 29 AA; 3464 MW; 58B785764E2623E3 CRC64;

Query Match 19.4%; Score 28; DB 1; Length 29;
 Best Local Similarity 55.6%; Pred. No. 5e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRR 10
 : : : : : : : : : : : : : : : :
 Db 18 DRKKRLEQ 26

RESULT 14
 GLUC_LAMFL
 ID GLUC_LAMFL STANDARD; PRT; 29 AA.
 AC Q9PRQ9;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucagon.
 OS Lampetra fluviatilis (River lamprey).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lampetra.
 OX NCBI_TaxID=7748;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Small intestine;
 RX MEDLINE=96108396; PubMed=8575665;
 RA Conlon J.M., Bondareva V., Rusakov Y., Plisetskaya E.M.,
 RA Mynarcik D.C., Whittaker J.;
 RT "Characterization of insulin, glucagon, and somatostatin from the
 RT river lamprey, Lampetra fluviatilis";
 RL Gen. Comp. Endocrinol. 100:96-105(1995).
 CC -1- FUNCTION: PROMOTES HYDROLYSIS OF GLYCOGEN AND LIPIDS, AND RAISES
 CC THE BLOOD SUGAR LEVEL.
 CC -1- INDUCTION: PRODUCED IN THE A CELLS OF THE ISLETS OF LANGERHANS
 CC IN RESPONSE TO A DROP IN BLOOD SUGAR CONCENTRATION.
 CC -1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
 DR HSP; P01274; IGCN.
 DR InterPro; IPR000532; Glucagon.
 DR Pfam; PF00123; hormone2; 1.
 DR PRINTS; PRO0275; GLUCAGON.
 DR SMART; SM00070; GLUCA; 1.

DR PROSITE; PS00260; GLUCAGON; 1.
 KW Glucagon family; Hormone.
 SQ SEQUENCE 29 AA; 3398 MW; 03A901D08C5EAB27 CRC64;

Query Match 19.1%; Score 27.5; DB 1; Length 29;
 Best Local Similarity 44.4%; Pred. No. 5.9e+02;
 Matches 8; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

OY 13 GSQT-DEFTLTSSLOAED 29
 : : : : : : : : : : : : : : : :
 Db 4 GSPTSDYSKYLDSKQAKD 21

RESULT 15
 NEUT_CAVPO
 ID NEUT_CAVPO STANDARD; PRT; 13 AA.
 AC P32560;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neurotensin (NT).
 GN NTS.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystriognath; Cavidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Small intestine;
 RX MEDLINE=86248085; PubMed=3087775;
 RA Shaw C., Thim L., Conlon J.M.;
 RT "[Ser7]neurotensin: isolation from guinea pig intestine";
 RL FEBS Lett. 202:187-192(1986).
 CC -1- FUNCTION: Smooth muscle-contracting peptide.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.
 DR PIR; A53608; A53608.
 KW Vasoactive.
 FT MOD_RES 1
 SQ SEQUENCE 13 AA; 1680 MW; 4C8314644CA115B3 CRC64;

Query Match 18.8%; Score 27; DB 1; Length 13;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRR 7
 : : : : : : : : : : : : : : : :
 Db 3 YENKSR 9

Search completed: July 15, 2002, 13:57:02
 Job time: 315 sec

•
•
•
•

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:52:12 ; Search time 39.48 Seconds

(without alignments)
127.073 Million cell updates/sec

Title: US-09-712-819A-9

Perfect score: 144
Sequence: 1 YERKKRRRRRRSGSGTDFLTITLISLQAE D 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 12865

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	29.9	26	13	Q90Y47
2	43	29.9	26	13	Q90V47
3	43	29.9	26	13	Q90V47
4	42	29.2	23	8	Q90V47
5	37	25.7	13	12	Q90V47
6	35	24.3	25	13	Q90V47
7	33	22.9	25	5	Q96252
8	33	22.9	27	13	Q90V47
9	33	22.9	27	13	Q90V47
10	32	22.2	29	6	Q9TSS7
11	31	21.5	23	15	Q85496
12	30	20.8	23	5	Q94781
13	30	20.8	28	10	Q42364
14	29	20.1	22	13	Q919P3
15	29	20.1	28	7	Q19358
16	29	20.1	29	15	Q91Q15

17	28	19.4	20	12	Q85617	Q85617 reovirus (t
18	28	19.4	21	12	Q90G30	Q91930 tt virus. o
19	28	19.4	21	12	Q90G26	Q91926 tt virus. o
20	28	19.4	21	12	Q90G23	Q91923 tt virus. o
21	28	19.4	22	12	Q85664	Q85664 reovirus (t
22	28	19.4	23	8	Q95770	Q95770 pinus mugo.
23	28	19.4	23	8	Q95758	Q95758 pinus sylve
24	28	19.4	26	11	Q99M13	Q99M13 mus musculu
25	28	19.4	27	5	Q95X71	Q95X71 caenorhabdi
26	28	19.4	28	12	Q9NM18	Q9NM18 tt virus. o
27	28	19.4	28	12	Q9NM10	Q9NM10 tt virus. o
28	28	19.4	29	15	P89816	P89816 human immun
29	28	19.4	29	15	P89821	P89821 human immun
30	27	18.8	15	13	Q9PRM3	Q9PRM3 galus gall
31	27	18.8	15	13	Q9R3F6	Q9R3F6 pseudomonas
32	27	18.8	22	8	Q95774	Q95774 abies alba
33	27	18.8	24	11	Q61946	Q61946 mus musculu
34	27	18.8	25	4	Q96PE8	Q96PE8 homo sapien
35	27	18.8	26	2	Q9R4S8	Q9R4S8 rhodobacter
36	27	18.8	26	8	P92646	P92646 lialis jica
37	27	18.8	26	15	Q04070	Q04070 human t-cel
38	27	18.8	27	12	Q90HB7	Q90HB7 hepatitis c
39	27	18.8	28	12	Q9NM19	Q9NM19 tt virus. o
40	26.5	18.4	15	2	Q47612	Q47612 escherichia
41	26.5	18.4	29	4	Q9UN87	Q9UN87 homo sapien
42	26.5	18.4	29	15	Q91Q16	Q91Q16 human immun
43	26	18.1	9	6	Q9T777	Q9T777 bos taurus
44	26	18.1	9	10	Q940K4	Q940K4 arabidopsis
45	26	18.1	13	7	Q19690	Q19690 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	26 AA.
Q90Y47	Q90Y47	Q90Y47	Q90Y47	Q90Y47
AC	Q90Y47	Q90Y47	Q90Y47	Q90Y47
DT	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	01-DEC-2001 (TREMBLrel. 19, Last sequence update)	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE	CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).	CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).	CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).	CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).
GN	CHD1-W.	CHD1-W.	CHD1-W.	CHD1-W.
OS	Cyanopsitta spixii x Ara macacana.	Cyanopsitta spixii x Ara macacana.	Cyanopsitta spixii x Ara macacana.	Cyanopsitta spixii x Ara macacana.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae.	Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae.	Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae.	Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae.
OX	NCBI_TaxID=157078;	NCBI_TaxID=157078;	NCBI_TaxID=157078;	NCBI_TaxID=157078;
RN	[1]	[1]	[1]	[1]
RP	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.
RA	Miyaki C.Y., Faria P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;	Miyaki C.Y., Faria P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;	Miyaki C.Y., Faria P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;	Miyaki C.Y., Faria P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;
RT	"The last Spix's Macaw and an Illiger's Macaw produced a hybrid."	"The last Spix's Macaw and an Illiger's Macaw produced a hybrid."	"The last Spix's Macaw and an Illiger's Macaw produced a hybrid."	"The last Spix's Macaw and an Illiger's Macaw produced a hybrid."
RL	Conserv. Genet. 2:53-55(2001).	Conserv. Genet. 2:53-55(2001).	Conserv. Genet. 2:53-55(2001).	Conserv. Genet. 2:53-55(2001).
DR	EMBL; AF276769; AAK96044.1; -.	EMBL; AF276769; AAK96044.1; -.	EMBL; AF276769; AAK96044.1; -.	EMBL; AF276769; AAK96044.1; -.
KW	Helicase.	Helicase.	Helicase.	Helicase.
FT	NON_TER	NON_TER	NON_TER	NON_TER
FT	NON_TER	NON_TER	NON_TER	NON_TER
SO	SEQUENCE	SEQUENCE	SEQUENCE	SEQUENCE
Query Match	26 AA;	26 AA;	26 AA;	26 AA;
Best Local Similarity	29.9%;	29.9%;	29.9%;	29.9%;
Mismatches	9;	9;	9;	9;
Conservative	2;	2;	2;	2;
Mismatches	5;	5;	5;	5;
Indels	0;	0;	0;	0;
Gaps	0;	0;	0;	0;
RESULT	2	2	2	2
Q90V47	Q90V47	Q90V47	Q90V47	Q90V47
ID	Q90V47	Q90V47	Q90V47	Q90V47
AC	Q90V47	Q90V47	Q90V47	Q90V47
DT	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)	01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
 DE CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).
 GN CHD1-W.
 OS Ara. maracana.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae; Ara.
 OX NCBI_TaxID=157077;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN=1, AND 2;
 RA Miyaki C.Y., Farla P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;
 RT "The last Spix's Macaw and an illiger's Macaw produced a hybrid.";
 RL Conserv. Genet. 2:53-55(2001).
 DR EMBL: AF276764; AAK96039.1; -;
 DR EMBL: AF276766; AAK96041.1; -;
 KW Helicase.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 26 AA; 2973 MW; F179E3EB84AC1297 CRC64;

Query Match 29.9%; Score 43; DB 13; Length 26;
 Best Local Similarity 56.2%; Pred. No. 10;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 ERKRRRRRRSGSGT 17
 ||: |||||:
 Db 7 ERHRSRRYSGSDD 22

RESULT 3
 ID 090V2 PRELIMINARY; PRT; 26 AA.
 AC 090V2;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).
 GN CHD1-W.
 OS Cyanopsitta spixii.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae;
 OC Cyanopsitta.
 OX NCBI_TaxID=51902;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA STRAIN=1, AND 2;
 RA Miyaki C.Y., Farla P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;
 RT "The last Spix's Macaw and an illiger's Macaw produced a hybrid.";
 RL Conserv. Genet. 2:53-55(2001).
 DR EMBL: AF276767; AAK96042.1; -;
 DR EMBL: AF276768; AAK96043.1; -;
 KW Helicase.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 26 AA; 2973 MW; F179E3EB84AC1297 CRC64;

Query Match 29.9%; Score 43; DB 13; Length 26;
 Best Local Similarity 56.2%; Pred. No. 10;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 ERKRRRRRRSGSGT 17
 ||: |||||:
 Db 7 ERHRSRRYSGSDD 22

RESULT 4
 ID 035223 PRELIMINARY; PRT; 23 AA.
 AC 035223;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
 DE RIBOSOMAL PROTEIN S19 (FRAGMENT).
 GN RPS19.
 OS Oenothera bertiana (Bertero's evening primrose).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Myrtales; Onagraceae; Oenothera.
 OX NCBI_TaxID=3950;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94250844; PubMed=8193306;
 RA Bock H., Brennicke A., Schuster W.;
 RT "Rps3 and rpl16 genes do not overlap in Oenothera mitochondria: GTC
 as a potential translation initiation codon in plant mitochondria?";
 RL Plant Mol. Biol. 24:811-818(1994).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92107682;
 RA Schuster W., Brennicke A.;
 RT "RNA editing makes mistakes in plant mitochondria: editing looses
 sense in transcripts of a rps19 pseudogene and in creating stop codons
 in cox1 and rps3 mRNAs of Oenothera.";
 RL Nucleic Acids Res. 19:6923-6928(1991).
 DR EMBL: X69140; CAA48893.1; -;
 KW Mitochondrion.
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 23 AA; 2764 MW; 08B5AB73D448C5CF CRC64;

Query Match 29.2%; Score 42; DB 8; Length 23;
 Best Local Similarity 53.3%; Pred. No. 13;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 YERKRRRRRRSGSG 15
 : ||: ||| |
 Db 3 FYRKRRTSRRTNGSG 17

RESULT 5
 ID 065331 PRELIMINARY; PRT; 13 AA.
 AC 065331;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DE 01-NOV-1998 (TReMBLrel. 08, last annotation update)
 DE BASIC PROTEIN (FRAGMENT).
 OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6;
 RX MEDLINE=90218045; PubMed=2109042;
 RA Hill-Perkins M.S., Possee R.D.;
 RT "A baculovirus expression vector derived from the basic protein
 promoter of Autographa californica nuclear polyhedrosis virus.";
 RL J. Gen. Virol. 71:971-976(1990).
 DR EMBL: D00864; BAA00738.1; -;
 FT NON_TER
 FT NON_TER
 SQ SEQUENCE 13 AA; 1626 MW; 6C75B50A749D1414 CRC64;

Query Match 25.7%; Score 37; DB 12; Length 13;
 Best Local Similarity 70.0%; Pred. No. 40;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 RORRRSGSGT 16
 ||: |||||:
 Db 4 RRRRRSTGT 13

```

RESULT 6
Q90Y57 PRELIMINARY; PRT; 25 AA.
AC Q90Y57;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE RIBOSOMAL PROTEIN L41.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Ictaluridae; Ictalurus.
OX NCBI_TaxId=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Patterson A.P., Karsl A., Liu Z.J.;
RT "Translational machinery of channel catfish: II. Complementary DNA and
RT Expression of the Complete Set of 47 60S Ribosomal Proteins."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF401598; AAC95170.1; -;
SQ SEQUENCE 25 AA; 3357 MW; 9AD629C69EDC029E CRC64;

```

```

Query Match 24.3%; Score 35; DB 13; Length 25;
Best Local Similarity 54.5%; Pred. No. 1.7e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 ERKRRQRRRS 12
Db 14 KRRRRKRRRS 24

```

```

RESULT 7
Q962S2 PRELIMINARY; PRT; 25 AA.
AC Q962S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE RIBOSOMAL PROTEIN L41.
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuidae; Noctuidae; Amphipyritinae; Spodoptera.
OX NCBI_TaxId=7108;
RN [1]
RP SEQUENCE FROM N.A.
RA Landais I., Ogliastro M., Mita K., Nohata J., Lopez-Ferber M.,
RA Duonot-Cerutti M., Fournier P., Devauchelle G.;
RT "Full-length ribosomal protein sequence from an EST library of
RT Spodoptera frugiperda cells (Stf9)."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF400204; AAK92176.1; -;
SQ SEQUENCE 25 AA; 3399 MW; 9AD629DD9EC039E CRC64;

```

```

Query Match 22.9%; Score 33; DB 5; Length 25;
Best Local Similarity 54.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 ERKRRQRRRS 12
Db 14 KRRRRKRRRS 24

```

```

RESULT 8
Q57542 PRELIMINARY; PRT; 27 AA.
AC Q57542;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

```

```

DE HOMEBOX PROTEIN LPHOXB (FRAGMENT).
OS Lampetra planeri (Brook lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxId=7750;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98358009; PubMed=9694633;
RA Sharmam A.C., Holland P.W.;
RT "Estimation of Hox gene cluster number in lampreys."
RL Int. J. Dev. Biol. 42:617-620(1998).
DR EMBL; AF044798; AAC03002.1; -;
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3098 MW; FE1EE5E971D0C1DA CRC64;

```

```

Query Match 22.9%; Score 33; DB 13; Length 27;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 3 RRRRRRRSGSGTD 17
Db 9 RRRRRRRRAAQRD 23

```

```

RESULT 9
Q57543 PRELIMINARY; PRT; 27 AA.
AC Q57543;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE HOMEBOX PROTEIN LPHOXB (FRAGMENT).
OS Lampetra planeri (Brook lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxId=7750;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98358009; PubMed=9694633;
RA Sharmam A.C., Holland P.W.;
RT "Estimation of Hox gene cluster number in lampreys."
RL Int. J. Dev. Biol. 42:617-620(1998).
DR EMBL; AF044799; AAC03003.1; -;
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3190 MW; B86EE73971D0DC06 CRC64;

```

```

Query Match 22.9%; Score 33; DB 13; Length 27;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 3 RRRRRRRSGSGTD 17
Db 9 RRRRRRRRAAQRD 23

```

```

RESULT 10
Q9TSS7 PRELIMINARY; PRT; 29 AA.
AC Q9TSS7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE WILM'S TUMOR PROTEIN 1 (FRAGMENT).
GN WTL.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxId=9685;
RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21015404; PubMed=1130975;
RA Brouillette J.A.; Andrew J.R., Venta P.J.;
RT "Estimate of nucleotide diversity in dogs with a pool-and-sequence
method."
RL Mamm. Genome 11:1079-1086(2000).
DR EMBL; AF201740; AAF19825.1; -
DR EMBL; AF201739; AAF19825.1; JOINED.
FT NON_TER 1 1
SQ SEQUENCE 29 AA; 3594 MW; FDC8E57A0BAE35DD CRC64;

Query Match 22.2%; Score 32; DB 6; Length 29;
Best Local Similarity 66.7%; Pred. No. 5, 8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 KRRRRRSG 13
DB 1 KRHRRTTG 9

RESULT 11
ID 085496 PRELIMINARY; PRT; 23 AA.
AC 085496;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Bovine leukemia virus (BLV).
OC Viruses; Retrovirdae; Retroviridae; Deltaretrovirus.
ON NCBI_TaxID=11901;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198886; PubMed=3033284;
RA Rice N.R., Simsek S.L., Dubois G.C., Showalter S.D., Gilden R.V.,
RA Stephens R.M.;
RT "Expression of the bovine leukemia virus X region in virus-infected
RT cells."
RL J. Virol. 61:1577-1585(1987).
DR EMBL; M16017; AAA87336.1; -.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2877 MW; F85E7CBED5440B08 CRC64;

Query Match 21.5%; Score 31; DB 15; Length 23;
Best Local Similarity 75.0%; Pred. No. 6, 4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 KRRRRRR 11
DB 3 KERRRRR 10

RESULT 12
ID 094781 PRELIMINARY; PRT; 23 AA.
AC 094781;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE HISTONE H2A (FRAGMENT).
OS Trypanosoma cruzi.
OC Eukaryota; Eulenzozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
ON NCBI_TaxID=5693;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RA Tanaka T., Tanaka M.;
RT "The parasite genome project: Trypanosoma cruzi."
RT Submitted (AUG-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; D87227; BAA13318.1; -.

FT NON_TER 1 1
SQ SEQUENCE 23 AA; 2790 MW; 12E9ED7592E52045 CRC64;

Query Match 20.8%; Score 30; DB 5; Length 23;
Best Local Similarity 66.7%; Pred. No. 9, 1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 KRRRRRR 11
DB 10 KRRRRRR 18

RESULT 13
ID 042364 PRELIMINARY; PRT; 28 AA.
AC 042364;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ALPHA-AMYLASE SUBFAMILY AMY1 PROTEIN (FRAGMENT).
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
ON NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92366494; PubMed=1502164;
RA Huang N., Stebbins G.L., Rodriguez R.L.;
RT "Classification and evolution of alpha-amylase genes in plants."
RL Proc. Natl. Acad. Sci. U.S.A. 89:7526-7530(1992).
DR EMBL; S42213; AAD13822.1; -.
DR HSSP; P04063; IAVA. 1
FT NON_TER 1 1
SQ SEQUENCE 28 AA; 3022 MW; 69EED0A4138FA43A1 CRC64;

Query Match 20.8%; Score 30; DB 10; Length 28;
Best Local Similarity 42.9%; Pred. No. 1, 1e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 12 SGGGDFTLTISL 25
DB 14 SGGGTFDFTFKGI 27

RESULT 14
ID 0919P3 PRELIMINARY; PRT; 22 AA.
AC 0919P3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CHROMO-HELICASE-DNA BINDING PROTEIN (FRAGMENT).
OS Ara militaris.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Palaeognathae; Psittacidae; Ara.
ON NCBI_TaxID=57237;
RN [1]
RP SEQUENCE FROM N.A.
RA Leal-Garza C.H., Bermudez-Humaran L.G., Jaramillo-Rangel G.,
RA Riojas-Valdes V., Montes de Oca-Puna R.;
RT "Molecular sexing of monomorphic endangered Ara birds."
RT Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF229847; AAF42915.1; -.
KM Helicase.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2703 MW; C8E13B3B1AA56873 CRC64;

Query Match 20.1%; Score 29; DB 13; Length 22;

Best Local Similarity 60.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERKKRRQRR 11
111:1:11
DB 4 ERKKRRKGR 13

RESULT 15

019358 PRELIMINARY; PRT; 28 AA.
AC 019358;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BOTA PROTEIN (FRAGMENT).
GN BOTA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052564; PubMed=1428011;
RA Ellis S.A., Braem K.A., Morrison W.I.;
RT "Transmembrane and cytoplasmic domain sequences demonstrate at least
two expressed bovine MHC class I loci."
RL Immunogenetics 37:49-56(1992).
DR EMBL; S47737; AAB23971.1; -.
FT NON_TER 1
SQ SEQUENCE 28 AA; 2852 MW; 78798D5C19C16CF4 CRC64;

Query Match 20.1%; Score 29; DB 7; Length 28;
Best Local Similarity 45.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 12 SSGTDPFTIT 22
111:1:111
DB 15 SAAGSDVSLTV 25

Search completed: July 15, 2002, 13:57:48
JOB time: 336 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:47:51 ; Search time 48.92 seconds
(without alignments)
65.845 Million cell updates/sec

Title: US-09-712-819a-9
Perfect score: 144
Sequence: 1 YERKKRRRRSGGTDTLTLSLQAE D 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 272466

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.GeneSeq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	41.0	16	20	AAV41874
2	59	41.0	16	22	AAU25389
3	59	41.0	16	22	AAU15733
4	56	38.9	18	17	AAAR6601
5	55	38.2	15	15	AAAR49577
6	55	38.2	15	18	AAAM3207
7	55	38.2	15	18	AAAM32916
8	55	38.2	15	18	AAAM26448
9	55	38.2	15	19	AAAW76153
10	55	38.2	16	18	AAW19818
11	55	38.2	22	22	AAE12486

12	55	38.2	22	22	AAE12487	Membrane transflin
13	55	38.2	24	15	AAAR49575	Residues 37-58GCG
14	55	38.2	24	18	AAAM31205	HIV-1 based transp
15	55	38.2	24	18	AAAM32914	HIV-1 tat protein
16	55	38.2	24	18	AAAM26446	HIV tat protein fr
17	55	38.2	24	19	AAAM76151	HIV Type I Tat pro
18	54	37.5	25	15	AAAR49567	Residues 38-62 HIV
19	54	37.5	25	18	AAAM31216	HIV-TAT protein tr
20	54	37.5	25	18	AAAM26456	Peptide of the Inv
21	54	37.5	25	19	AAAM76162	HIV Type I Tat pro
22	54	37.5	26	11	AAAR03434	AAS 37-62 of HIV t
23	54	37.5	26	11	AAAR02238	Tat-dm37-62, 46Ala
24	54	37.5	26	11	AAAR02238	Tat-dm37-62, 40Ala
25	54	37.5	26	11	AAAR04719	Tat-dm37-62, 39Ala
26	54	37.5	26	11	AAAR02240	Tat-dm37-62, 38Ala
27	54	37.5	26	11	AAAR02241	Tat-dm37-62, 37Ala
28	54	37.5	26	13	AAAR30086	pH-dependent membr
29	54	37.5	26	13	AAAR49558	Residues 47-72 HIV
30	54	37.5	26	15	AAAR49558	HIV-TAT protein tr
31	54	37.5	26	18	AAAM31211	Peptide of the Inv
32	54	37.5	26	18	AAAM26451	HIV Type I Tat pro
33	54	37.5	26	19	AAAM76157	HIV transactivator
34	54	37.5	26	21	AAV96570	Human p75NTR depen
35	54	37.5	28	20	AAV33518	Preferred signal s
36	54	37.5	29	19	AAAM63399	Signal sequence fo
37	54	37.5	29	21	AAV53820	Membrane transport
38	54	37.5	29	22	AAE11950	RNA binding peptid
39	53	36.8	13	21	AAV69771	HIV-1 tat derived
40	53	36.8	13	22	AAU06108	Peptide associated
41	52	36.8	25	22	AAAB67911	Tat-CD3 carboxyl t
42	52	36.1	21	22	AAAB57547	Beta-catenin modul
43	52	36.1	21	22	AAAB57957	
44	52	36.1	21	22	AAAB03933	
45	51.5	35.8	22	21	AAAB03933	

ALIGNMENTS

RESULT 1
ID AAV41874 standard; Peptide: 16 AA.
AC AAV41874;
DT 09-DEC-1999 (first entry)
XX
XX
DE Rheumatoid arthritis diagnostic protein isoform peptide #25.
KW Human; rheumatoid arthritis; RA; diagnosis; RPI; RADP; detection;
KW Rheumatoid arthritis diagnostic feature; ERPI; synovial fluid;
KW Rheumatoid arthritis diagnostic protein isoform; screening;
KW expression reference protein isoform; prognosis.
XX
XX Homo sapiens.
XX
XX WO947925-A2.
XX
XX PD 23-SEP-1999.
XX
XX PF 15-MAR-1999; 99WO-GB00763.
XX
XX PR 13-MAR-1998; 98GB-0005477.
XX
XX PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
XX PI Parekh RB, Patel TP, Townsend RR;
XX WPI: 1999-571871/48.
XX
XX DR
XX PT
XX PT Diagnosis of human rheumatoid arthritis by two-dimensional
XX electrophoresis -

PS Disclosure: Page 18; 157pp; English.
XX
CC A method has been developed for the diagnosis of human rheumatoid
CC arthritis (RA) using two-dimensional electrophoresis to generate a
CC two-dimensional array of features. The method can be used for screening,
CC diagnosis and prognosis of RA in a subject or for monitoring the effect
CC of an anti-RA drug or therapy administered to a subject. The method
CC comprises: (a) analysing a sample of serum or plasma and optionally
CC synovial fluid by two-dimensional electrophoresis, to generate a two-
CC dimensional array of features; (b) identifying at least one chosen
CC feature whose relative abundance correlates with the presence or absence
CC of RA; and (c) comparing the abundance of each chosen feature in the
CC sample with the abundance of that chosen feature in serum or plasma from
CC one or more persons without RA, where the relative abundance of the
CC chosen feature or features in the sample indicates the presence or
CC absence of RA in the subject. The method can also be used in clinical
CC studies for testing drugs for therapy of RA, for purification of RA-
CC diagnostic protein isoforms (RPIs), and for production of antibodies to
CC RPIs. The RA-diagnostic feature (RADF) proteins can be used to identify
CC RA drugs. Nucleic acid encoding RADFs can be used in gene therapy
CC protocols. AA141844 to AA142100 represent RPI peptides, AA142101 to
CC AA142103 represent expression reference protein isoform peptides and
CC AA225066 to AA225068 represent degenerate probes for RPIs, which are all
XX used in the exemplification of the present invention.
XX
SQ Sequence 16 AA;

Query Match 41.0%; Score 59; DB 20; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SGGSGTDFLTITIS 23
|||||
DB 4 sggsgtdftltis 15

RESULT 2
AAU25389 standard; Peptide; 16 AA.
XX
AC AAU25389;
XX
DT 18-DEC-2001 (first entry)
XX
DE Schizophrenia-Associated Protein Isoform (SPI) peptide #618.
XX
KW Schizophrenia-associated protein isoform; SPI; SPI-206; SPI-238; SPI-240;
KW neuroleptic; gene therapy; cerebrospinal fluid; serum; plasma.
XX
OS Homo sapiens.
XX
PN MO200162785-A2.
XX
PD 30-AUG-2001.
XX
PE 23-FEB-2001; 2001WO-GB00792.
XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAG, Parekh RB, Rohlf C, Terrett JA, Tyson KL;
XX
DR WPI; 2001-570624/64.
XX
PT New schizophrenia associated protein isoforms and encoding nucleic acid
PT molecules, useful for treatment, diagnosis and prognosis of
PT schizophrenia and screening for potential drugs for treatment and new
PT drug targets -
XX

PS Disclosure: Page 41; 148pp; English.
XX
CC The sequence represents a schizophrenia-associated protein isoform (SPI).
CC These protein isoforms, e.g. SPI-206, SPI-238 and SPI-240 are detectable
CC in cerebrospinal fluid, serum or plasma and are useful markers of
CC schizophrenia. The sequences can be used for treatment and diagnosis of
CC schizophrenia, screening, prognosis, monitoring the results of therapy,
CC identifying patients most likely to respond to a particular therapy and
CC identification of new targets for drug treatment. SPI DNA is useful as a
CC nucleic acid probe to detect the presence of nucleic acids or SPIs.
XX
SQ Sequence 16 AA;

Query Match 41.0%; Score 59; DB 22; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 SGGSGTDFLTITIS 23
|||||
DB 4 sggsgtdftltis 15

RESULT 3
AAU15733 standard; Peptide; 16 AA.
XX
AC AAU15733;
XX
DT 24-OCT-2001 (first entry)
XX
DE Schizophrenia-associated isoform peptide #618.
XX
KW Schizophrenia; neuroleptic; diagnostic; neuropsychiatric disorder;
KW neurological disorder; neuropathy.
XX
OS Homo sapiens.
XX
PN MO200163293-A2.
XX
PD 30-AUG-2001.
XX
PE 23-FEB-2001; 2001WO-GB00783.
XX
PR 24-FEB-2000; 2000GB-0004415.
PR 28-NOV-2000; 2000US-0750395.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Herath HMAG, Parekh RB, Rohlf C;
XX
DR WPI; 2001-502868/55.
XX
PT Diagnosing and monitoring Schizophrenia by detecting the presence of
PT Schizophrenia Associated Features and Schizophrenia Associated Protein
PT Isoforms in samples of cerebrospinal fluid -
XX
PS Claim 6; Page 41; 160pp; English.
XX
CC The invention relates to methods and compositions for screening,
CC diagnosis and prognosis of Schizophrenia. The method involves detecting
CC the presence of Schizophrenia (SCH) Associated Features (SFS) and SCH
CC Associated Protein Isoforms (SPIs) in samples, e.g. by electrophoresis,
CC immunoassay or hybridisation assay, for diagnosing and monitoring SCH,
CC studying the effectiveness of treatments and for identifying potential
CC therapeutic agents. The method is used for (1) screening or diagnosis of
CC SCH and the relative abundance of at least 1 chosen feature correlates
CC with the presence or absence of SCH; and (2) monitoring the effect of
CC therapy administered to a subject with SCH and the relative abundance of
CC at least 1 chosen feature which correlates with the severity of SCH.
CC The expression and activity of the SFS, SPIs and related molecules
CC (e.g. secondary messengers) are studied to diagnose SCH, monitor the
CC progress of the disorder and the effectiveness of treatment and as

CC targets to identify and produce potential therapeutic agents for the
 CC treatment of SCH. The paucity of detectable neurologic defects
 CC distinguishes neuropsychiatric disorders such as SCH from neurological
 CC disorders, where manifestations of anatomical and biochemical changes
 CC have been identified in many cases. Consequently the identification and
 CC characterisation of cellular and/or molecular causative defects and
 CC neuropathies are necessary for improved treatment of neuropsychiatric
 CC disorders. AAU1514-AAU15762 represent the amino acid sequences of
 CC schizophrenia-associated isoforms used in the method of the invention.
 XX

SO Sequence 16 AA:

Query Match 41.0%; Score 59; DB 22; Length 16;
 Best Local Similarity 100.0%; Pred. No. 0.089;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 SGSGTDFLTITS 23
 |||||
 DB 4 sgsgtdfltlts 15

RESULT 4

AA86601
 ID AAR86601 standard; peptide; 18 AA.

AC AAR86601;

DT 28-JUN-1996 (first entry)

DE HIV TAT cellular uptake region Immunogen #1.

KW Immunogen; cellular uptake region; transactivating protein; TAT; HIV;
 KM viraemia; antibody.

OS Human immunodeficiency virus type 1.

XX WO953199-A1.

PD 30-NOV-1995.

PE 16-MAY-1995; 95WO-US06077.

PR 23-MAY-1994; 94US-0247991.

PA (IMMU-) IMMUNOLOGY RES INST INC.

PI Culler MD, Goldstein G, Shenbagamurthi P;

DR WPI; 1996-049298/05.

PT Immunogen containing cellular uptake region of viral TAT protein -
 PT induces high antibody titre against TAT, partic. for preventing or
 PT treating HIV infection

PS Claim 6; Page 45; 54pp; English.

CC The sequences given in AAR86601-19 are immunogens derived from the
 CC cellular uptake region of the transactivating protein, TAT of HIV.
 CC These peptides are based on the region comprising amino acids 46-63
 CC (see AAR86625) of the HIV TAT protein. These peptides are used to
 CC protect against infection by HIV. They may also be used to reduce
 CC viraemia in already infected patients. These immunogens may be used
 CC to identify, or generate, antibodies specific for TAT.

SO Sequence 18 AA:

Query Match 38.9%; Score 56; DB 17; Length 18;
 Best Local Similarity 68.8%; Pred. No. 0.27;
 Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YERKKRRRRSGSGT 16

DB 2 YRKKRRRRRRRPPGS 17
 | |||||

RESULT 5

AA849577
 ID AAR49577 standard; peptide; 15 AA.

AC AAR49577;

DT 23-SEP-1994 (first entry)

DE Residues tat47-58GGC HIV tat protein.

KW HIV; tat; transport; moiety; conjugate; cargo molecule;
 KM cytoplasmic delivery; nuclear delivery; cysteine-rich region;
 KN transactivation; disulphide aggregation.

OS Human immunodeficiency virus type 1.

XX WO9404686-A.

PD 03-MAR-1994.

PE 19-AUG-1993; 93WO-US07833.

PR 21-AUG-1992; 92US-0934375.

PA (BIOJ) BIOGEN INC.

PI Barsoun JG, Fawell SE, Peplinsky RB;

DR WPI; 1994-083202/10.

PT Transport of cargo moieties into cells - using genetic fusions of
 PT chemical conjugates comprising a portion of HIV tat protein as
 PT transport peptide

PS Disclosure; Page 87; 153pp; English.

CC The sequences given in AAR49574-77 are fragments of the HIV tat protein
 CC which were used as transport moieties in the conjugate of the
 CC invention. These conjugates are used as cargo molecules for the
 CC efficient cytoplasmic and nuclear delivery of biologically active
 CC non-tat proteins, nucleic acids and other molecules which are not
 CC inherently capable of entering a target cell. The tat basic region
 CC amino acid sequence carries cargo molecules by covalent attachment.
 CC The reduced size of the transport peptides minimises interference
 CC with the biological activity of the cargo molecule. By virtue of
 CC the absence of the cysteine-rich region (residues 22-36) of the tat
 CC protein, problems of spurious transactivation and disulphide
 CC aggregation are solved. The reduced transport peptide size also
 CC enhances uptake efficiency.

SO Sequence 15 AA:

Query Match 38.2%; Score 55; DB 15; Length 15;
 Best Local Similarity 84.6%; Pred. No. 0.31;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRRRSG 13
 | |||||
 DB 1 YRKKRRRRRRPPG 13

RESULT 6

AAW31207
 ID AAW31207 standard; peptide; 15 AA.

AC AAW31207;

DT 20-MAR-1998 (first entry)

```
XX HIV-1 based transport polypeptide Tat47-58GCC.
DE Human immunodeficiency virus; HIV Type 1; Tat protein;
XX cargo molecules; intracellular delivery; fusion protein;
KM therapeutic; prophylactic; diagnostic; transport polypeptide.
XX Synthetic.
OS Human immunodeficiency virus type 1.
XX
XX Key Location/Qualifiers
FH Region 3..11 /label= basic_region
FT MISC-difference 13..15 /label= spacer_residues
FT
XX
XX US5674980-A.
XX
XX PD 07-OCT-1997.
XX
XX PF 21-DEC-1989; 89US-0454450.
XX
XX PR 28-APR-1994; 94US-0235403.
XX 21-DEC-1989; 89US-0454450.
XX PR 02-JAN-1991; 91US-0636662.
XX 21-AUG-1992; 92US-0934375.
XX PR 19-AUG-1992; 93WO-US07833.
XX PR 24-NOV-1993; 93US-0158015.
XX PR 25-MAY-1995; 95US-0450098.
XX
XX PA (BARS/) BARSOUM J G.
XX PA (FAME/) FAWELL S E.
XX PA (FRAN/) FRANKEL A.
XX PA (PABO/) PABO.
XX PA (PEPI/) PEPINSKY R B.
XX
XX PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX WPI; 1997-502388/46.
XX
XX PT Fusion proteins containing truncated HIV tat sequences - useful for
XX intracellular delivery of viral repressor proteins
XX
XX PS Example 1: Column 55; 77pp; English.
XX
XX CC This sequence represents the transport polypeptide Tat47-58GCC which can
XX be used in a novel method for the delivery of biologically active cargo
XX molecules into the cytoplasm and nuclei of eukaryotic cells. The tat
XX protein from immunodeficiency virus (e.g. HIV-1, HIV-2, SIV) is readily
XX taken up into cells when present extracellularly. This polypeptide
XX has been constructed from the HIV-1 tat protein (see AAW31202) by
XX removing the cysteine-rich region and the carboxy-terminal domain while
XX retaining the basic region. Cargo molecules could be covalently linked
XX to the protein to form a fusion protein without the existing problems of
XX spurious trans-activation and disulphide aggregation. The reduced size
XX of transport polypeptides also minimises interference with the
XX biological activity of the cargo molecule. This is applicable for
XX therapeutic, prophylactic or diagnostic intracellular delivery of small
XX molecules and macromolecules e.g. proteins, nucleic acids and
XX polysaccharides.
XX
XX SQ Sequence 15 AA;
```

```
Query Match 38.2%; Score 55; DB 18; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.31;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy 1 YERKKRRRRSG 13
Db 1 YGRKKRRRRSG 13
```

```
RESULT 7
AAW32916
ID AAW32916 standard; peptide; 15 AA.
XX
XX AC AAW32916;
XX
XX DT 19-JAN-1998 (first entry)
XX
XX DE HIV-1 tat protein residues 47-58 plus Gly-Gly-Cys linker.
XX
XX KW Chemical conjugate; transport polypeptide; tat protein;
XX nucleic acid; delivery; cytoplasm; cell nucleus; therapy;
XX prophylaxis; diagnosis; spurious transactivation; HIV-1;
XX disulphide aggregation.
XX
XX OS Human immunodeficiency virus type 1.
XX
XX PN US5670617-A.
XX
XX PD 23-SEP-1997.
XX
XX PF 21-DEC-1989; 89US-0454450.
XX
XX PR 28-APR-1994; 94US-0235403.
XX 21-DEC-1989; 89US-0454450.
XX PR 02-JAN-1991; 91US-0636662.
XX 21-AUG-1992; 92US-0934375.
XX PR 19-AUG-1992; 93WO-US07833.
XX PR 24-NOV-1993; 93US-0158015.
XX PR 25-MAY-1995; 95US-0450246.
XX
XX PA (BARS/) BARSOUM J G.
XX PA (FAME/) FAWELL S E.
XX PA (FRAN/) FRANKEL A.
XX PA (PABO/) PABO.
XX PA (PEPI/) PEPINSKY R B.
XX
XX PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX WPI; 1997-479523/44.
XX
XX PT Conjugate for intracellular delivery - comprising transport moiety
XX having amino acids 49-57 of human immunodeficiency virus tat protein
XX and nucleic acid cargo moiety
XX
XX PS Disclosure: Columns 57-58; 77pp; English.
XX
XX CC A novel covalently linked chemical conjugate, comprises a transport
XX polypeptide moiety consisting of at least residues 49-57 of human
XX immunodeficiency virus (HIV) tat protein, but not residues 22-36
XX and 73-86, e.g. the present peptide, and a cargo moiety comprising
XX a single or double stranded nucleic acid. The conjugate can be used
XX to deliver cargo moieties into the cytoplasm and nuclei of cells
XX for therapeutic, prophylactic and diagnostic applications. The HIV
XX tat protein is readily taken up into cells and the cell nucleus.
XX The reduced size of the transport polypeptides minimises
XX interference with the biological activity of the cargo molecule.
XX In addition, by virtue of the absence of the cysteine rich region
XX of the HIV tat protein, the transport polypeptides solve the
XX potential problems of spurious transactivation and disulphide
XX aggregation.
XX
XX SQ Sequence 15 AA;
```

```
Query Match 38.2%; Score 55; DB 18; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.31;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Oy 1 YERKKRRRRSG 13
Db 1 YGRKKRRRRSG 13
```

RESULT 8
AAW6448
ID AAW6448 standard; Protein: 15 AA.
XX
AC AAW6448;
XX
DT 16-DEC-1997 (first entry)
XX
DE HIV tat protein fragment (aa47-58GCC).
XX
KW HIV; tat protein; transport protein; cargo delivery.
XX
OS Chimeric human immunodeficiency virus type 1;
OS Chimeric synthetic.
XX
PN US5652122-A.
PD 29-JUL-1997.
XX
PF 21-DEC-1989; 89US-0454450.
XX
PR 28-APR-1994; 94US-0235403.
PR 21-DEC-1989; 89US-0454450.
PR 02-JAN-1991; 91US-0636662.
PR 21-AUG-1992; 92US-0934375.
PR 19-AUG-1993; 93WO-US07833.
PR 24-NOV-1993; 93US-0158015.
PR 25-MAY-1995; 95US-0450257.
XX
PA (BARS/) BARSOUM J G.
PA (FAME/) FAWELL S E.
PA (FRAN/) FRANKEL A.
PA (PABO/) PABO C.
PA (PEPI/) PEPIISKY R B.
PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX
XX WPI: 1997-392943/36.
DR
XX
XX New DNA constructs for transporting molecules to cells - encode a
PT fusion protein comprising a modified HIV tat protein and a
PT carboxy-terminal cargo moiety
XX
PS Disclosure: Column 57-58; 76pp; English.
XX
CC This polypeptide comprises amino acids 47-58 of the HIV tat
CC protein (see also AAW6443), plus a Gly-Gly-Cys C-terminal extension.
CC Novel transport proteins comprise modified HIV tat covalently
CC attached to a cargo molecule (see AAW6436-42). The modified tat
CC protein (see AAW6444-49) is characterised by the presence of the tat
CC basic region (amino acids 47-59), and the absence of the Cys-rich
CC region (amino acids 22-36) and the exon 2-encoded C-terminal region
CC (amino acids 73-86). These modifications solve the potential
CC problems of spurious trans-activation and disulphide aggregation,
CC while the reduced size of the transport protein minimises
CC interference with the biological activity of the cargo molecule.
CC DNA molecules that encode the modified tat fusion proteins are
CC claimed and can be used to deliver polypeptides or nucleic acids to
CC the cytoplasm or nuclei of cells in vivo or in vitro.
XX
SQ Sequence 15 AA:

Query Match 38.2%; Score 55; DB 18; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.31;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRRRRRSG 13
| | | | | | | | | |
Db 1 ygrfkkrrrrrrpg 13

RESULT 9
AAW76153
ID AAW76153 standard; protein; 15 AA.
XX
AC AAW76153;
XX
DT 24-NOV-1998 (first entry)
XX
DE HIV Type I TAT protein fragment tat47-58GCC.
XX
KW TAT protein; cargo molecule; therapy; diagnosis; transport protein;
KW fusion protein; human papillomavirus E2 repressor; target cell.
XX
OS Human immunodeficiency virus type 1.
OS
XX
PN US5804604-A.
PD 08-SEP-1998.
XX
PF 25-MAY-1995; 95US-0450236.
XX
PR 28-APR-1994; 94US-0235403.
PR 21-DEC-1989; 89US-0454450.
PR 02-JAN-1991; 91US-0636662.
PR 19-AUG-1993; 93WO-US07833.
PR 24-NOV-1993; 93US-0158015.
PR 25-MAY-1995; 95US-0450236.
XX
PA (BIOJ) BIOGEN INC.
XX
PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX
XX WPI: 1998-505702/43.
DR
XX
XX HIV tat-derived transport fusion proteins - used to deliver
PT biological active molecules e.g. peptide(s) or nucleic acids,
PT specifically into cytoplasm or nuclei of cells
XX
PS Disclosure: Column 57-58; 83pp; English.
XX
XX
XX This sequence represents a human immunodeficiency virus (HIV) Type I
CC TAT protein fragment, tat47-58GCC which is used in a method for the
CC delivery of biologically active cargo molecules into the cytoplasm and
CC nuclei of cells, for therapeutic, prophylactic or diagnostic purposes.
CC This is accomplished by the presence of a small, basic section of tat
CC transport protein of HIV. This is used as it is this protein which is
CC involved to cause human cells in culture to take up HIV. The method
CC involves the use of a cargo moiety in combination with a transport
CC moiety usually in the form of a fusion protein. The cargo moiety is a
CC human papillomavirus E2 repressor that retains its biological activity
CC after delivery into a target cell and where the transport moiety is
CC one of following HIV tat protein fragments (a) aa 47-58, (b) aa 47-72,
CC (c) 38-72, (d) aa 38-58, (e) aa 37-58, (f) aa 1-21 and 38-72,
CC (g) aa 47-62 or aa 38-62. The proteins allow delivery of specific
CC peptides into cells at high concentrations due to use of existing
CC transporters. Previous methods of delivery include bombardment and
CC transfection, which only allow a fraction of the cell population to be
CC infected and can additionally damage cells as they cause physical opening
CC of the cell walls/membranes to allow entry.
XX
SQ Sequence 15 AA:

Query Match 38.2%; Score 55; DB 19; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.31;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRRRRRSG 13
| | | | | | | | | |
Db 1 ygrfkkrrrrrrpg 13

RESULT 10

AAW19818
 ID AAW19818 standard; Peptide: 16 AA.
 XX
 AC AAW19818;
 XX
 DT 26-JAN-1998 (first entry)
 XX
 DE Universal transfer vector sequence of chimeric coat protein.
 XX
 KM Adenovirus; vector: coat protein; gene therapy; gene transfer;
 KM human; cancer; autoimmune disease; heart disease; infection;
 KM universal transfer vector; PBSS 75-100 pgs(tat).
 XX
 OS Synthetic.
 XX
 PN WO9720051-A2.
 XX
 PD 05-JUN-1997.
 XX
 PF 27-NOV-1996; 96WO-US19150.
 XX
 PR 21-AUG-1996; 96US-0701124.
 PR 28-NOV-1995; 95US-0563368.
 PR 21-AUG-1996; 96US-0700846.
 XX
 PA (GENV-) GENVEC INC.
 XX
 PI Brough DE, Kovasdi I, Wickham TJ;
 XX
 DR WPI: 1997-310606/28.
 DR N-PSDB; AAT70272.
 XX
 PT Adenoviral vectors containing chimeric coat protein - bind and enter
 PT cells more efficiently, useful for gene therapy of e.g. cancer,
 PT autoimmune diseases, etc.
 XX
 PS Claim 7; Page 59-60; 121pp; English.
 XX
 CC This peptide can be used as a universal transfer vector (UTV)
 CC and/or spacer sequence in novel chimeric adenovirus coat proteins
 CC (CPs). The peptide is encoded by a gene fragment (see AAT70272) of
 CC transfer plasmid PBSS 75-100 pgs(tat). Claimed UTVs/spacers are
 CC given in AAW19810-11, AAW19813-25, AAW19827, AAW19829, AAW19831-32 and
 CC AAW19834-43). Claimed CPs, such as fibre proteins, differ from wild-
 CC type CPs by the introduction of the UTV and/or spacer at or near
 CC the C-terminus. This imparts on the chimeric CP the ability to
 CC bind to and enter cells by means of a novel cell surface binding
 CC site. Adenoviral vectors comprising the chimeric CP are able to
 CC enter cells more efficiently than vectors comprising wild-type CP,
 CC especially at lower m.o.i. They are especially useful for gene
 CC therapy of e.g. cancer, genetic disorders, pathogenic infections,
 CC heart disease or autoimmune diseases.
 CC
 XX
 SQ Sequence 16 AA;

Query Match 38.2%; Score 55; DB 18; Length 16;
 Best Local Similarity 78.6%; Pred. No. 0.33;
 Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSGS 14
 I | | | | | | | | | |
 DB 3 ygrkkrrqrrrrss 16

RESULT 11
 AAE12486
 ID AAE12486 standard; peptide: 22 AA.
 XX
 AC AAE12486;
 XX
 DT 03-JAN-2002 (first entry)
 XX

DE Membrane transiting antiviral peptide bTAT-9.
 XX
 KM Membrane transiting peptide; virucide; antiviral; Herpes Simplex Virus;
 KM HSV; HIV; Human Immunodeficiency Virus; CMV; cytomegalovirus.
 XX
 OS Unidentified.
 XX
 FH Key
 FT Modified-site 1 Location/Qualifiers
 FT /note= "Biotin-aminohexanoyl Tyr"
 XX
 PN WO200157072-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 06-FEB-2001; 2001WO-US03813.
 XX
 PR 07-FEB-2000; 2000US-180823P.
 PR 22-FEB-2000; 2000US-184057P.
 XX
 PA (MISC) WISCONSIN ALUMNI RES FOUND.
 XX
 PI Brandt C, Bultmann H;
 XX
 DR WPI: 2001-638840/73.
 XX
 PT Peptides comprising membrane transiting peptides useful for treating or
 PT preventing a virus infection, e.g., human immunodeficiency virus,
 PT herpes simplex virus and cytomegalovirus -
 XX
 PS Claim 7; Page 15; 43pp; English.
 XX
 CC The invention relates to peptides comprising membrane transiting
 CC peptides with antiviral properties. The peptides are useful for
 CC treating or preventing a virus infection in a warm blooded animal,
 CC e.g., enveloped viruses such as human immunodeficiency virus (HIV),
 CC herpes simplex virus (HSV), cytomegalovirus (CMV) and non-enveloped
 CC viruses. Preferably, the peptides are useful for treating or preventing
 CC infections from one or more HSVs. The antiviral peptides are used for
 CC treating viral infections of the skin or part of the oral or genital
 CC cavity. The present sequence is membrane transiting antiviral peptide.
 CC
 XX
 SQ Sequence 22 AA;

Query Match 38.2%; Score 55; DB 22; Length 22;
 Best Local Similarity 84.6%; Pred. No. 0.46;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSG 13
 I | | | | | | | | | |
 DB 1 ygrkkrrqrrrrp 13

RESULT 12
 AAE12487
 ID AAE12487 standard; peptide: 22 AA.
 XX
 AC AAE12487;
 XX
 DT 03-JAN-2002 (first entry)
 XX
 DE Membrane transiting antiviral peptide bTAT-9x.
 XX
 KM Membrane transiting peptide; virucide; antiviral; Herpes Simplex Virus;
 KM HSV; HIV; Human Immunodeficiency Virus; CMV; cytomegalovirus.
 XX
 OS Unidentified.
 XX
 FH Key
 FT Modified-site 1 Location/Qualifiers
 FT /note= "Biotin-aminohexanoyl Tyr"

```

PN  W0200157072-A2.
XX
XX  09-AUG-2001.
PD
XX
XX  06-FEB-2001; 2001WO-US03813.
PF
XX
XX  07-FEB-2000; 2000US-180823P.
PR
XX  22-FEB-2000; 2000US-184057P.
PA
XX  (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
XX  Brandt C, Bullmann H;
PI
XX
XX  WPI: 2001-638840/73.
DR
XX
XX  Peptides comprising membrane transiting peptides useful for treating or
PT  preventing a virus infection, e.g., human immunodeficiency virus,
PT  herpes simplex virus and cytomegalovirus -
XX
XX  Claim 7; Page 15; 43pp; English.
PS
XX
XX  The invention relates to peptides comprising membrane transiting
CC  peptides with antiviral properties. The peptides are useful for
CC  treating or preventing a virus infection in a warm blooded animal,
CC  e.g., enveloped viruses such as human immunodeficiency virus (HIV),
CC  herpes simplex virus (HSV), cytomegalovirus (CMV) and non-enveloped
CC  viruses. Preferably, the peptides are useful for treating or preventing
CC  infections from one or more HSVs. The antiviral peptides are used for
CC  treating viral infections of the skin or part of the oral or genital
CC  cavity. The present sequence is membrane transiting antiviral peptide.
CC
XX
SQ  Sequence 22 AA:

Query Match 38.2%; Score 55; DB 22; Length 22;
Best Local Similarity 84.6%; Pred. No. 0.46;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSG 13
   1 | | | | | | | | | |
DB 1 YRKRRRRRRRTPG 13

RESULT 13
AAR49575
ID  AAR49575 standard; peptide; 24 AA.
XX
XX  AAR49575;
AC
XX
XX  23-SEP-1994 (first entry)
DT
XX
XX  Residues 37-58GGC HIV tat protein.
DE
XX
XX  HIV; tat; transport; moiety; conjugate; cargo molecule;
KW  cytoplasmic delivery; nuclear delivery; cysteine-rich region;
KW  transactivation; disulphide aggregation.
XX
XX  Human immunodeficiency virus type 1.
OS
XX
XX  W09404686-A.
PN
XX
XX  03-MAR-1994.
PD
XX
XX  19-AUG-1993; 93WO-US07833.
PF
XX
XX  21-AUG-1992; 92US-0934375.
PR
XX
XX  (BIOJ ) BIOGEN INC.
PA
XX
XX  Barsoun JG, Fawell SE, Pepinsky RB.
XX
XX  WPI: 1994-083202/10.
DR
XX

```

```

PT  Transport of cargo moieties into cells - using genetic fusions of
PT  chemical conjugates comprising a portion of HIV tat protein as
PT  transport peptide
XX
XX  Disclosure; Page 86; 153pp; English.
PS
XX
XX  The sequences given in AAR49574-77 are fragments of the HIV tat protein
CC  which were used as transport moieties in the conjugate of the
CC  invention. These conjugates are used as cargo molecules for the
CC  efficient cytoplasmic and nuclear delivery of biologically active
CC  non-tat proteins, nucleic acids and other molecules which are not
CC  inherently capable of entering a target cell. The tat basic region
CC  amino acid sequence carries cargo molecules by covalent attachment.
CC  The reduced size of the transport peptides minimises interference
CC  with the biological activity of the cargo molecule. By virtue of
CC  the absence of the cysteine-rich region (residues 22-36) of the tat
CC  protein, problems of spurious transactivation and disulphide
CC  aggregation are solved. The reduced transport peptide size also
CC  enhances uptake efficiency.
CC
XX
XX  Sequence 24 AA:
SQ

Query Match 38.2%; Score 55; DB 15; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.5;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSG 13
   1 | | | | | | | | | |
DB 10 YRKRRRRRRRTPG 22

RESULT 14
AAW31205
ID  AAW31205 standard; peptide; 24 AA.
XX
XX  AAW31205;
AC
XX
XX  20-MAR-1998 (first entry)
DT
XX
XX  HIV-1 based transport polypeptide Tat38-58GGC.
DE
XX
XX  Human immunodeficiency virus; HIV Type 1; Tat protein;
KW  cargo molecules; intracellular delivery; fusion protein;
KW  therapeutic; prophylactic; diagnostic; transport polypeptide.
XX
XX  Synthetic.
OS
XX
XX  Human immunodeficiency virus type 1.
XX
XX  Key Location/Qualifiers
FH 12..20
FT Region /label= basic_domain
FT Misc-difference 22..24 /note= "spacer residues"
FT
XX
XX  US5674980-A.
PN
XX
XX  07-OCT-1997.
PD
XX
XX  21-DEC-1989; 89US-0454450.
PF
XX
XX  28-APR-1994; 94US-0235403.
PR
XX  21-DEC-1989; 89US-0454450.
PR  02-JAN-1991; 91US-0636662.
PR  21-AUG-1992; 92US-0934375.
PR  19-AUG-1993; 93WO-US07833.
PR  24-NOV-1993; 93US-0158015.
PR  25-MAY-1995; 95US-0450098.
XX
XX  (BARS/) BARSOUM J G.
XX  (FAWE/) FAWELL S E.
XX  (FRAN/) FRANKEL A.
XX  (PABO/) PABO.
PA

```

PA (PEPI/) PEPINSKY R B.
 XX
 PI Barsom JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 XX
 XX WPI; 1997-502388/46.
 DR
 XX
 PT Fusion proteins containing truncated HIV tat sequences - useful for
 PT intracellular delivery of viral repressor proteins
 XX
 PS Example 1; Column 55; 77pp; English.
 XX
 CC This sequence represents the transport polypeptide Tat38-586GC which can
 CC be used in a novel method for the delivery of biologically active cargo
 CC molecules into the cytoplasm and nuclei of eukaryotic cells. The tat
 CC protein from immunodeficiency virus (e.g. HIV-1, HIV-2, SIV) is readily
 CC taken up into cells when present extracellularly. This polypeptide
 CC has been constructed from the HIV-1 tat protein (see AMW1202) by
 CC removing the cysteine-rich region and the carboxy-terminal domain while
 CC retaining the basic region. Cargo molecules could be covalently linked
 CC to the protein to form a fusion protein without the existing problems of
 CC spurious trans-activation and disulphide aggregation. The reduced size
 CC of transport polypeptides also minimises interference with the
 CC biological activity of the cargo molecule. This is applicable for
 CC therapeutic, prophylactic or diagnostic intracellular delivery of small
 CC molecules and macromolecules e.g. proteins, nucleic acids and
 CC polysaccharides.
 CC
 SQ Sequence 24 AA;

Query Match 38.2%; Score 55; DB 18; Length 24;
 Best Local Similarity 84.6%; Pred. No. 0.5;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YERKKRRRRSG 13
 I | | | | | | | | | |
 DB 10 YGRKKRRGRRPG 22

RESULT 15
 AAW32914
 ID AAW32914 standard; peptide; 24 AA.
 XX
 AC AAW32914;
 XX
 DT 19-JAN-1998 (first entry)
 XX
 DE HIV-1 tat protein residues 38-58 plus Gly-Gly-Cys linker.
 XX
 KW Chemical conjugate; transport polypeptide; tat protein;
 KW nucleic acid; delivery; cytoplasm; cell nucleus; therapy;
 KW prophylaxis; diagnosis; spurious transactivation; HIV-1;
 KW disulphide aggregation.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN US5670617-A.
 XX
 PD 23-SEP-1997.
 XX
 PF 21-DEC-1989; 89US-0454450.
 XX
 XX 28-APR-1994; 94US-0235403.
 PR 21-DEC-1989; 89US-0454450.
 PR 02-JAN-1991; 91US-0636662.
 PR 21-AUG-1992; 92US-0934375.
 PR 19-AUG-1993; 93WO-US07833.
 PR 24-NOV-1993; 93US-0158015.
 PR 25-MAY-1995; 95US-0450246.
 XX
 PA (BARS/) BARSOM J G.
 PA (FAWE/) FAWELL S E.
 PA (FRAN/) FRANKEL A.

PA (PABO/) PABO C.
 PA (PEPI/) PEPINSKY R B.
 XX
 XX Barsom JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 PI
 XX WPI; 1997-479523/44.
 DR
 XX
 PT Conjugate for intracellular delivery - comprising transport moiety
 PT having amino acids 49-57 of human immunodeficiency virus tat protein
 PT and nucleic acid cargo moiety
 XX
 PS Disclosure; Columns 55-56; 77pp; English.
 XX
 CC A novel covalently linked chemical conjugate, comprises a transport
 CC polypeptide moiety consisting of at least residues 49-57 of human
 CC immunodeficiency virus (HIV) tat protein, but not residues 22-36
 CC and 73-86, e.g. the present peptide, and a cargo moiety comprising
 CC a single or double stranded nucleic acid. The conjugate can be used
 CC to deliver cargo moieties into the cytoplasm and nuclei of cells
 CC for therapeutic, prophylactic and diagnostic applications. The HIV
 CC tat protein is readily taken up into cells and the cell nucleus.
 CC The reduced size of the transport polypeptides minimises
 CC interference with the biological activity of the cargo molecule.
 CC In addition, by virtue of the absence of the cysteine rich region
 CC of the HIV tat protein, the transport polypeptides solve the
 CC potential problems of spurious transactivation and disulphide
 CC aggregation.
 CC
 SQ Sequence 24 AA;

Query Match 38.2%; Score 55; DB 18; Length 24;
 Best Local Similarity 84.6%; Pred. No. 0.5;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 YERKKRRRRSG 13
 I | | | | | | | | | |
 DB 10 YGRKKRRGRRPG 22

Search completed: July 15, 2002, 13:51:41
 Job time: 230 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:49:36 ; Search time 21.58 Seconds
(without alignments)
32.824 Million cell updates/sec

Title: US-09-712-819A-9
Perfect score: 144
Sequence: 1 YERKKRRRRSGSGTDTLTLSQAED 29

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 142465

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79	54.9	25	5	PCT-US91-02942-42
2	75	52.1	25	5	PCT-US91-02942-26
3	60	41.7	13	3	US-08-783-853A-60
4	56	38.9	18	5	PCT-US95-06077-3
5	56	38.9	26	5	PCT-US95-06077-24
6	55	38.2	15	1	US-08-450-257-6
7	55	38.2	15	1	US-08-450-246-6
8	55	38.2	15	1	US-08-450-098-6
9	55	38.2	15	1	US-08-451-233-6
10	55	38.2	15	1	US-08-450-236-6
11	55	38.2	15	4	US-08-235-403-6
12	55	38.2	24	1	US-08-450-257-4
13	55	38.2	24	1	US-08-450-246-4
14	55	38.2	24	1	US-08-450-098-4
15	55	38.2	24	1	US-08-451-233-4
16	55	38.2	24	1	US-08-450-236-4
17	55	38.2	24	4	US-08-235-403-4
18	54	37.5	25	1	US-08-450-257-53
19	54	37.5	25	1	US-08-450-246-53
20	54	37.5	25	1	US-08-450-098-53
21	54	37.5	25	1	US-08-451-233-53
22	54	37.5	25	1	US-08-450-236-53
23	54	37.5	25	4	US-08-235-403-53
24	54	37.5	26	1	US-08-450-257-48
25	54	37.5	26	1	US-08-450-246-48
26	54	37.5	26	1	US-08-450-098-48
27	54	37.5	26	1	US-08-451-233-48

28	54	37.5	26	1	US-08-450-236-48	Sequence 48, Appl
29	54	37.5	26	4	US-08-235-403-48	Sequence 48, Appl
30	54	37.5	26	6	516933-2	Patent No. 516933
31	54	37.5	28	4	US-09-041-886-47	Sequence 47, Appl
32	54	37.5	29	2	US-08-928-958-9	Sequence 9, Appl
33	54	37.5	29	2	US-09-072-429-9	Sequence 9, Appl
34	54	37.5	29	4	US-09-411-706-5	Sequence 5, Appl
35	51	35.4	14	5	PCT-US95-06077-4	Sequence 4, Appl
36	51	35.4	14	5	PCT-US95-06077-10	Sequence 10, Appl
37	51	35.4	14	5	PCT-US95-06077-12	Sequence 12, Appl
38	51	35.4	25	4	US-09-041-886-40	Sequence 40, Appl
39	50.5	35.1	25	4	US-09-041-886-37	Sequence 37, Appl
40	50.5	35.1	28	4	US-09-041-886-38	Sequence 38, Appl
41	50.5	35.1	28	4	US-09-041-886-39	Sequence 39, Appl
42	50.5	35.1	28	4	US-09-041-886-53	Sequence 53, Appl
43	50.5	35.1	28	4	US-09-041-886-54	Sequence 54, Appl
44	50.5	35.1	28	4	US-09-041-886-54	Sequence 54, Appl
45	50	34.7	11	2	US-08-706-741B-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
PCT-US91-02942-42
; Sequence 42, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: ATWAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02942
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009549.8
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAM L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 1011.0586600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; PCT-US91-02942-42

Query Match 54.9%; Score 79; DB 5; Length 25;
Best Local Similarity 88.9%; Pred. No. 42e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 12 SSGSDFTLTLSQAED 29
||||| :|

Db 2 SSGSTDFTLTISLQPPD 19

RESULT 2

PCT-US91-02942-26

Sequence 26, Application PC/TUS9102942

GENERAL INFORMATION:

APPLICANT: ROTHLEIN, ROBERT

APPLICANT: ADAIR, JOHN R

APPLICANT: ATTMAL, DILBERT S

TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY

NUMBER OF SEQUENCES: 102

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox

STREET: 1225 Connecticut Ave. NW Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02942

FILING DATE: 19910429

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9009549.8

FILING DATE: 27-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: FOX, SAM L

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 1011.0586600

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 466-0800

TELEFAX: (202) 833-8716

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US91-02942-26

Query Match 52.1%; Score 75; DB 5; Length 25;

Best Local Similarity 83.3%; Pred. No. 0.00016;

Matches 15; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 12 SSGSTDFTLTISLQPPD 29

Db 2 SSGSTDFTLTISLQPPD 19

RESULT 3

US-08-783-853A-60

Sequence 60, Application US/08783853A

PATENT NO. 6005091

GENERAL INFORMATION:

APPLICANT: Blackburn, Michael

APPLICANT: Church, William

APPLICANT: Gross, Mitchell

APPLICANT: Feuerstein, Gloria

APPLICANT: Nichols, Andrew

APPLICANT: Padlan, Eduardo

APPLICANT: Patel, Arunbhai

APPLICANT: Sylvester, Daniel

TITLE OF INVENTION: ANTICOAGULANT AGENTS USEFUL IN TREATMENT

TITLE OF INVENTION: OF THROMBOSIS

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:

ADDRESSEE: SmithKline Beecham Corporation

STREET: 709 Swedeland Road

CITY: King of Prussia

STATE: PA

COUNTRY: USA

ZIP: 19406

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/783, 853A

FILING DATE: 16-JAN-1997

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/029, 119

FILING DATE: 24-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Baumeister, Kirk

REGISTRATION NUMBER: 33,833

REFERENCE/DOCKET NUMBER: P50438

TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-270-5096

TELEFAX:

TELEX:

INFORMATION FOR SEQ ID NO: 60:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

US-08-783-853A-60

Query Match 41.7%; Score 60; DB 3; Length 13;

Best Local Similarity 92.3%; Pred. No. 0.011;

Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 12 SSGSTDFTLTIS 24

Db 1 SSGSTDFTLTIS 13

RESULT 4

PCT-US95-06077-3

Sequence 3, Application PC/TUS9506077

GENERAL INFORMATION:

APPLICANT: Immunobiology Research, Institute Inc.

TITLE OF INVENTION: Vaccine Interdiction of Extracellular

TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus

TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar

TITLE OF INVENTION: Intercellular Transactivating Strategies

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, PO Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06077

FILED DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/247,991
FILING DATE: 23-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: IRI44PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-06077-3

Query Match 38.9%; Score 56; DB 5; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.057;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YEKRRRORRRSGSGT 16
1 | | | | | | | | | | : 1 :
Db 2 YGKRRRORRRAPGS 17

RESULT 5
PCT-US95-06077-24
Sequence 24, Application PC/TUS9506077
GENERAL INFORMATION:
APPLICANT: Immunobiology Research, Institute Inc.
TITLE OF INVENTION: Vaccine Interdiction of Extracellular
TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus
TITLE OF INVENTION: and Other Chronically Infecting Vitrues Employing Similar
TITLE OF INVENTION: Intercellular Transactivating Strategies
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06077
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/247,991
FILING DATE: 23-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: IRI44PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-06077-24

Query Match 38.9%; Score 56; DB 5; Length 26;
Best Local Similarity 68.8%; Pred. No. 0.086;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YEKRRRORRRSGSGT 16
1 | | | | | | | | | | : 1 :
Db 2 YGKRRRORRRAPGS 17

RESULT 6
US-08-450-257-6
Sequence 6, Application US/08450257
Patent No. 5652122
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: BARO, Carl
APPLICANT: BARSOUD, James G.
APPLICANT: PAMEL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,257
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-450-257-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.065;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRRRSG 13
|1111111111|
DB 1 YGKKRRRRRRPG 13

RESULT 7
US-08-450-246-6
; Sequence 6, Application US/08450246
; Patent No. 5670617
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPIISKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,246
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9090
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-450-246-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.065;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRRRSG 13
|1111111111|

DB 1 YGKKRRRRRRPG 13

RESULT 8
US-08-450-098-6
; Sequence 6, Application US/08450098
; Patent No. 5674980
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPIISKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,098
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9090
; TELEFAX: (212) 596-9090
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-450-098-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.065;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRRRSG 13
|1111111111|
DB 1 YGKKRRRRRRPG 13

RESULT 9

US-08-451/233-6
; Sequence 6, Application US/08451233
; Patent No. 5747641
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,233
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-451-233-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.065;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YEKRRRRRRRG 13
| | | | | | | | | | |
Db 1 YEKRRRRRRRG 13

RESULT 10
US-08-450-236-6
; Sequence 6, Application US/08450236
; Patent No. 5804604
; GENERAL INFORMATION:

APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,236
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-450-236-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.065;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YEKRRRRRRRG 13
| | | | | | | | | | |
Db 1 YEKRRRRRRRG 13

RESULT 11
US-08-235-403-6
; Sequence 6, Application US/08235403
; Patent No. 6316003
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.

```

? APPLICANT: PEPINSKY, R. B.
? TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
? NUMBER OF SEQUENCES: 69
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FISH & NEAVE
? STREET: 1251 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10020
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/235,403
? FILING DATE:
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/934,375
? FILING DATE: 21-AUG-1992
? APPLICATION NUMBER: US 07/098,766
? FILING DATE: 28-JUL-1993
? APPLICATION NUMBER: PCT/US93/07833
? FILING DATE: 19-AUG-1993
? APPLICATION NUMBER: US 07/454,450
? FILING DATE: 21-DEC-1989
? APPLICATION NUMBER: US 07/636,662
? FILING DATE: 02-JAN-1991
? APPLICATION NUMBER: US 08/158,015
? FILING DATE: 24-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 27,794
? REFERENCE/DOCKET NUMBER: B170 CIP 2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 596-9000
? TELEFAX: (212) 596-9090
? TELEX: 14-8367
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 15 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
?
? US-08-235-403-6
?
? Query Match 38.2%; Score 55; DB 4; Length 15;
? Best Local Similarity 84.6%; Pred. No. 0.065;
? Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
?
? QY 1 YERKKRRRRRSG 13
? 1 YERKKRRRRRSG 13
? DB 1 YGRRKKRRRRRPG 13
?
? RESULT 12
? US-08-450-257-4
? Sequence 4, Application US/08450257
? Patent No. 5652122
? GENERAL INFORMATION:
? APPLICANT: FRANKEL, Alan
? APPLICANT: PABO, Carl
? APPLICANT: BARSOUM, James G.
? APPLICANT: FAWELL, Stephen E.
? APPLICANT: PEPINSKY, R. B.
? TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
? NUMBER OF SEQUENCES: 69
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FISH & NEAVE
? STREET: 1251 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10020

```

```

? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10020
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/450,257
? FILING DATE: 25-MAY-1995
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/235,403
? FILING DATE: 28-APR-1994
? APPLICATION NUMBER: US 07/934,375
? FILING DATE: 21-AUG-1992
? APPLICATION NUMBER: US 07/098,766
? FILING DATE: 28-JUL-1993
? APPLICATION NUMBER: PCT/US93/07833
? FILING DATE: 19-AUG-1993
? APPLICATION NUMBER: US 07/454,450
? FILING DATE: 21-DEC-1989
? APPLICATION NUMBER: US 07/636,662
? FILING DATE: 02-JAN-1991
? APPLICATION NUMBER: US 08/158,015
? FILING DATE: 24-NOV-1993
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 27,794
? REFERENCE/DOCKET NUMBER: B170 CIP 2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 596-9000
? TELEFAX: (212) 596-9090
? TELEX: 14-8367
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 24 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
?
? US-08-450-257-4
?
? Query Match 38.2%; Score 55; DB 1; Length 24;
? Best Local Similarity 84.6%; Pred. No. 0.11;
? Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
?
? QY 1 YERKKRRRRRSG 13
? 1 YERKKRRRRRSG 13
? DB 10 YGRRKKRRRRRPG 22
?
? RESULT 13
? US-08-450-246-4
? Sequence 4, Application US/08450246
? Patent No. 5670617
? GENERAL INFORMATION:
? APPLICANT: FRANKEL, Alan
? APPLICANT: PABO, Carl
? APPLICANT: BARSOUM, James G.
? APPLICANT: FAWELL, Stephen E.
? APPLICANT: PEPINSKY, R. B.
? TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
? NUMBER OF SEQUENCES: 69
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FISH & NEAVE
? STREET: 1251 Avenue of the Americas
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10020

```

```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,098
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLGY: linear
MOLECULE TYPE: peptide
US-08-450-098-4

Query Match      38.2% Score 55; DB 1; Length 24;
Best Local Similarity 84.6% Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY      1 YEKRRRRRRRG 13
      | | | | | | | | |
Db      10 YGRKKRRRRRPG 22

RESULT 15
US-08-451-233-4
; Sequence 4, Application US/08451233
; Patent No. 5747641
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPIINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/451,233
; FILING DATE: 25-MAY-1995

```

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: HALEY JR., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ. ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-451-233-4

Query Match 38.2%; Score 55; DB 1; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YEKRRRRRRSG 13
| | | | | | | | | |
DB 10 YGKKRRRRRRPG 22

Search completed: July 15, 2002, 13:52:08
Job time: 152 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:52:41 ; Search time 25.22 Seconds

(without alignments)
110.491 Million cell updates/sec

Title: US-09-712-819A-10

Perfect score: 144

Sequence: 1 YERKKRRRRSGSGTDAAPASSIQAEED 29

Scoring table: BL0SUM62

Gapop 10.0 , Gapept 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 5929

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	29.2	23	2	S43765
2	39	27.1	29	2	S08555
3	38	26.4	27	1	SRAPAS
4	35	24.3	21	2	PN0082
5	35	24.3	25	1	JC4278
6	35	24.3	25	2	J01617
7	35	24.3	25	2	JC4685
8	35	24.3	25	2	S38425
9	35	24.3	25	2	T06233
10	35	24.3	25	2	T49214
11	34	23.6	25	2	A60741
12	34	23.6	27	1	SRAPC
13	33	22.9	15	2	A38304
14	33	22.9	25	2	T38719
15	32	22.2	28	2	S26254
16	31	21.5	25	1	R6BY48
17	29	20.1	20	2	S09720
18	29	20.1	22	2	C64330
19	29	20.1	27	2	PC2316
20	29	20.1	29	2	S14040
21	28	19.4	10	2	PH0946
22	28	19.4	12	2	A53524
23	28	19.4	15	2	A49480
24	28	19.4	16	2	A35552
25	28	19.4	19	2	A33443
26	28	19.4	21	2	PC4443
27	28	19.4	23	2	A36150
28	28	19.4	23	2	A48185
29	28	19.4	25	2	S07250

30	28	19.4	28	2	T37143	hypothetical prote
31	28	19.4	28	2	S68643	nicotinic acetylch
32	27.5	19.1	23	2	S47192	T-cell receptor J-
33	27	18.8	13	2	A53608	neurotensin - guin
34	27	18.8	17	2	E23734	insulin-like growt
35	27	18.8	18	2	PS0387	platelet-derived g
36	27	18.8	20	2	S38291	30K allergen - vel
37	27	18.8	23	2	S41390	p7 protein - human
38	27	18.8	27	2	S09504	trans-activating t
39	27	18.8	29	1	TNLJBR	trans-activating t
40	27	18.8	19	2	S35924	T-cell receptor ga
41	26.5	18.4	19	2	S12268	Oa-2 antigen - mou
42	26	18.1	10	2	E60787	sperm-activating p
43	26	18.1	13	1	UNBO	neurotensin - bovi
44	26	18.1	20	2	A61327	trypsin (EC 3.4.21
45	26	18.1	20	2	A25335	myosin-light-chain

ALIGNMENTS

RESULT 1

S43765 ribosomal protein S19 - evening primrose mitochondrion (fragment)

C:Species: mitochondrion Oenothera villarica (evening primrose)

C:Date: 10-Dec-1994 #sequence #revision 12-May-1995 #text_change 13-Aug-1999

C:Accession: S43765; S30542

R:Bock, H.; Brennick, A.; Schuster, W.

Plant Mol. Biol. 24, 811-818, 1994

A:Title: Rps3 and rpl16 genes do not overlap in Oenothera mitochondria: GTG as a pote

A:Reference number: S43765; MUID:94250844

A:Accession: S43765

A:Molecule type: DNA

A:Residues: 1-23 <BOC>

A:Cross-references: EMBL:X69140; NID:G13188; PIDN:CAA4893.1; PID:G13189

A:Note: the source is designated as Oenothera berteriana

C:Genetics:

A:Gene: rps19

A:Genome: mitochondrion

C:Superfamily: Escherichia coli ribosomal protein S19

C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match 29.2%; Score 42; DB 2; Length 23;

Best Local Similarity 53.3%; Pred. No. 41;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 YERKKRRRRSGSG 15

Db 3 FTRKKRRSRITNGS 17

RESULT 2

S08555 ribosomal protein L15 [validated] - Halobacterium salinarum (fragment)

N:Alternate names: ribosomal protein HL16

C:Species: Halobacterium salinarum

C:Date: 31-Dec-1990 #sequence #revision 31-Dec-1990 #text_change 21-Jul-2000

C:Accession: S08555

R:Matheson, A.T.; Yaguchi, M.; Christensen, P.; Rollin, C.F.; Hasnain, S.

Can. J. Biochem. Cell Biol. 62, 426-433, 1984

A:Title: Purification, properties, and N-terminal amino acid sequence of certain 50S

A:Reference number: S07437; MUID:84282108

A:Accession: S08555

A:Molecule type: protein

A:Residues: 1-29 <MAT>

A:Note: the source is designated as Halobacterium cutirubrum

A:Note: the protein is designated as ribosomal protein L16

C:Superfamily: rat ribosomal protein L27a

C:Keywords: protein biosynthesis; ribosome

Query Match 27.1%; Score 39; DB 2; Length 29;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

OY 4 KRRROR--RRSGGT 16
| | | | | |
Db 3 KRRRORGRTHGGGT 17

RESULT 3

SRAPS
N:Alternate names: stellate sturgeon
C:Species: Acipenser stellatus (stellate sturgeon)
C:Date: 28-Feb-1980 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997
C:Accession: A02666
R:Yulikova, E.P.; Rybin, V.K.; Silaev, A.B.
Bioorg. Khim. 5, 5-10, 1979
A:Title: The primary structure of stellin A.
A:Reference number: A02666
A:Accession: A02666
A:Molecule type: protein
A:Residues: 1-27 <YU>
A:Note: article in Russian with English abstract
C:Superfamily: protamine Y2
C:Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match 26.4%; Score 38; DB 1; Length 27;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 4 KRRRORRRSG 13
| | | | | |
Db 13 KRRRORRRRHG 22

RESULT 4

PN0082
sperm chromatin protein I2-2 - Argentinian shortfin squid (fragment)
C:Species: Illex argentinus (Argentinian shortfin squid)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PN0082
R:Osadchuk, L.A.; Levina, N.B.; Telezhinskaya, I.N.; Khrapunov, S.N.; Berdyshev, G.D.; A
Bioorg. Khim. 16, 448-455, 1990
A:Title: Primary structure of main nuclear protein from headleg mollusk Illex argentinus
A:Reference number: PN0081; MUID:90329055
A:Accession: PN0082
A:Molecule type: protein
A:Residues: 1-21 <OSA>
A:Experimental source: sperm
C:Superfamily: sperm histone
C:Keywords: sperm

Query Match 24.3%; Score 35; DB 2; Length 21;
Best Local Similarity 58.3%; Pred. No. 3e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 RRRRRRRSGS 14
| | | | | |
Db 1 RRRRRRRRRRS 12

RESULT 5

JC4278
ribosomal protein L41, cytosolic [similarity] - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 01-Sep-2000
C:Accession: JC4278
R:Chan, Y.L.; Olivera, J.; Wool, I.G.
Biochem. Biophys. Res. Commun. 214, 810-818, 1995
A:Title: The primary structures of rat ribosomal proteins L4 and L41.
A:Reference number: JC4277; MUID:96024571

A:Accession: JC4278
A:Molecule type: mRNA
A:Residues: 1-25 <CHA>
A:Cross-references: EMBL:X82550; NID:g575383; PIDN:CAA57899.1; PID:g575384
C:Superfamily: rat ribosomal protein L41
C:Keywords: ribosome

Query Match 24.3%; Score 35; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
| | | | | |
Db 14 KRRRRRRORS 24

RESULT 6

J01617
ribosomal protein L41.e - human
N:Alternate names: HG12 protein; ribosomal protein YL41
C:Species: Homo sapiens (man)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: J01617; JC5659; S31691
R:Klaudiny, J.; von der Kammer, H.; Scheit, K.H.
Biochem. Biophys. Res. Commun. 187, 901-906, 1992
A:Title: Characterization by cDNA cloning of the mRNA of a highly basic human protein
A:Reference number: J01617; MUID:92412140
A:Accession: J01617
A:Molecule type: mRNA
A:Residues: 1-25 <KLA>
A:Cross-references: EMBL:Z12962; NID:g36135; PIDN:CAA78306.1; PID:g36136
R:Lee, J.H.; Kim, J.M.; Kim, M.S.; Lee, Y.T.; Marshak, D.R.; Bae, Y.S.
Biochem. Biophys. Res. Commun. 238, 462-467, 1997
A:Title: The highly basic ribosomal protein L41 interacts with the beta subunit of pr
A:Reference number: JC5659; MUID:97446005
A:Accession: JC5659
A:Molecule type: mRNA
A:Residues: 1-25 <LEE>
A:Cross-references: EMBL:Z12962; NID:g36135; PIDN:CAA78306.1; PID:g36136
C:Comment: This protein stimulates phosphorylation of the beta chain of DNA topoisom
C:Superfamily: rat ribosomal protein L41

Query Match 24.3%; Score 35; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
| | | | | |
Db 14 KRRRRRRORS 24

RESULT 7

JC4685
ribosomal protein L41 - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 18-Aug-2000
C:Accession: JC4685
R:Starkay, C.R.; Menon, R.P.; Prabhu, S.; Levy, L.S.
Biochem. Biophys. Res. Commun. 220, 648-652, 1996
A:Title: Primary sequence and evolutionary conservation of ribosomal protein genes fr
A:Reference number: JC4685; MUID:96183078
A:Accession: JC4685
A:Molecule type: mRNA
A:Residues: 1-25 <STA>
A:Cross-references: GB:U22229; NID:g950108; PIDN:AAB01667.1; PID:g950109
A:Experimental source: thymic lymphoma
C:Genetics:
A:Gene: rPL41
C:Superfamily: rat ribosomal protein L41

Query Match 24.3% Score 35; DB 2; Length 25;
Best Local Similarity 54.5% Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
:||:|:|:|
DB 14 KKKRRKMRORS 24

RESULT 8
S38425

ribosomal protein GlA1 - upland cotton

C:Species: Gossypium hirsutum (upland cotton)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Aug-2000

C:Accession: S38425

R:Turley, R.B.; Ferguson, D.L.; Meredith, W.R.

submitted to the EMBL data library, October 1993

A:Reference number: S38425

A:Accession: S38425

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-25 <TUR>

A:Cross-references: EMBL:X75423; NID:9407800; PIDN:CA53175.1; PID:9407801

C:Superfamily: rat ribosomal protein L41

Query Match 24.3% Score 35; DB 2; Length 25;
Best Local Similarity 54.5% Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
:||:|:|:|
DB 14 KKKRRKMRORS 24

RESULT 9
T06233

ribosomal protein L41 - soybean (fragment)

C:Species: Glycine max (soybean)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 18-Aug-2000

C:Accession: T06233

R:Mahalingam, R.; Knap, H.T.

submitted to the EMBL data library, February 1998

A:Reference number: Z15555

A:Accession: T06233

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-25 <MAH>

A:Cross-references: EMBL:AF047051; NID:92905777; PIDN:AAC03557.1; PID:92905778

A:Experimental source: cultivar Essex; root

C:Superfamily: rat ribosomal protein L41

C:Keywords: protein biosynthesis; ribosome

Query Match 24.3% Score 35; DB 2; Length 25;
Best Local Similarity 54.5% Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
:||:|:|:|
DB 14 KKKRRKMRORS 24

RESULT 10
T49214

ribosomal protein GlA1-like - Arabidopsis thaliana

N:Alternate names: protein F27K19.200

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000

C:Accession: T49214

R:Benes, V.; Wurmbech, E.; Drzonek, H.; Ansoerge, W.; Mewes, H.W.; Rudd, S.; Lemcke, K.;

submitted to the Protein Sequence Database, April 2000

A:Reference number: Z25014

A:Accession: T49214
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-25 <BEN>
A:Cross-references: EMBL:AL163832; GSPDB:GN00061; ATSP:F27K19.200
A:Experimental source: cultivar Columbia; BAC clone F27K19

C:Genetics:

A:Gene: ATSP:F27K19.200

A:Map position: 3

A:Introns: 4/3

C:Superfamily: rat ribosomal protein L41

Query Match 24.3% Score 35; DB 2; Length 25;
Best Local Similarity 54.5% Pred. No. 3.5e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKKRRRRRS 12
:||:|:|:|
DB 14 KKKRRKMRORS 24

RESULT 11
A60741

insulin-like growth factor-binding protein He39L - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993

C:Accession: A60741

R:Forbes, B.; Ballard, F.J.; Wallace, J.C.

J. Endocrinol. 126, 497-506, 1990

A:Title: An insulin-like growth factor-binding protein purified from medium condition

A:Reference number: A60741; MIMD:91011238

A:Accession: A60741

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-25 <FOR>

Query Match 23.6% Score 34; DB 2; Length 25;
Best Local Similarity 35.3% Pred. No. 4.7e+02;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 13 GSGTDPAPASLQAPD 29
| | | | |
DB 6 GQGVQAGXPGGXVEED 22

RESULT 12
SRAPC

protamine B - Russian sturgeon

N:Alternate names: sturine

C:Species: Acipenser gueldenstaedti (Russian sturgeon)

C:Date: 12-Aug-1981 #sequence_revision 12-Aug-1981 #text_change 16-Feb-1997

C:Accession: A02665

R:Yulikova, E.P.; Evseenko, L.K.; Baratova, L.A.; Belyanova, L.P.; Ryblin, V.K.; Sllae

Biorg. Khim. 2, 1613-1617, 1976

A:Reference number: A02665

A:Accession: A02665

A:Molecule type: protein

A:Residues: 1-27 <YUL>

C:Superfamily: protamine Y2

C:Keywords: chromosomal protein; DNA binding; spermatogenesis

Query Match 23.6% Score 34; DB 1; Length 27;
Best Local Similarity 60.0% Pred. No. 5e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KKKRRRRRG 13
:||:|:|:|
DB 11 QRRRRRRRHG 20

RESULT 13

A38304
heterogeneous ribonuclear particle protein A1 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 31-Oct-1997
C:Accession: A38304
R:Kumar, A.; Casas-Pineb, J.R.; Luneau, C.J.; Karpel, R.L.; Merrill, B.M.; Williams, K.F.
J. Biol. Chem. 265, 17094-17100, 1990
A:Title: Mammalian heterogeneous nuclear ribonucleoprotein A1. Nucleic acid binding prop
A:Reference number: A38304; MUID:91009136
A:Accession: A38304
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-15 <RUM>
C:Superfamily: helix-destabilizing protein; ribonucleoprotein repeat homology

Query Match 22.9%; Score 33; DB 2; Length 15;
Best Local Similarity 87.5%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 8 ORRRSGSG 15
DB 8 ORRRSGSG 15

RESULT 14

T38719
ribosomal protein L41 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T38719; T38731; T39176
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z21807
A:Accession: T38719
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25 <MUR>
A:Cross-references: EMBL:Z69369; PIDN:CAB40187.1; GSPDB:GN00066
A:Experimental source: strain 972h-; cosmid c3p10
R:Genies, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21797
A:Accession: T38731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25 <GEN>

A:Cross-references: EMBL:Z99167; PIDN:CAB40152.1; GSPDB:GN00066; SPDB:SPAC366.13c
A:Experimental source: strain 972h-; cosmid c3c6
R:Bothe, G.; Pohl, T.; McDougall, R.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21833
A:Accession: T39176
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25 <BOT>

A:Cross-references: EMBL:AL109738; PIDN:CAB52162.1; GSPDB:GN00066; SPDB:SPAC8F11.01c
A:Experimental source: strain 972h-; cosmid c8F11
C:Genetics:
A:Gene: SPAC31F10.18c; SPAC366.13c; SPAC3F10.18c; SPAC8F11.01c
A:Map position: 1
A:introns: 24/2
C:Superfamily: rat ribosomal protein L41

Query Match 22.9%; Score 33; DB 2; Length 25;
Best Local Similarity 54.5%; Pred. No. 6.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKRRRQRRS 12
DB 14 KRRRRPRARS 24

RESULT 15

S26254
rel protein - chicken
C:Species: Gallus gallus (chicken)
C>Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 06-Dec-1996
C:Accession: S26254
R:Capobianco, A.J.; Gilmore, T.D.
Oncogene 6, 2203-2210, 1991
A:Title: Repression of the chicken c-rel promoter by vrel in chicken embryo fibroblas
A:Reference number: S26254; MUID:92115319
A:Accession: S26254
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-28 <CAP>
A:Cross-references: EMBL:X59588

Query Match 22.2%; Score 32; DB 2; Length 28;
Best Local Similarity 36.4%; Pred. No. 9.3e+02;
Matches 8; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

OY 1 YERKKRRRRSGSGTDTATAPA 22
DB 5 FVRGAARSPRRGGPSSDAELSA 26

Search completed: July 15, 2002, 13:52:41
Job time: 160 sec

"The primary structure of stellin A.:"

RL Bloerg. Khlm. 5:5-10(1979).
RN [2]
RP SEQUENCE (STELLIN C).
RX MEDLINE-94019256; Pubmed-1339950;
RA Rybin V.R., Revina L.P., Baratova L.A.;
RT "Comparison of amino acid sequences of sturgeon triptamines using
RT proteinases from *Acipenser stellatus* gonads as an example.";
RL Mol. Biol. (Mosk) 26:300-306(1992).
CC -1- FUNCTION: PROTAGENES SUBSTRATE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: Testis.
DR PIR: A02666; SRPAS.
KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
FT VARIANT 1 1 MISSING (IN STELLIN C).
SQ SEQUENCE 27 AA; 3532 MW; 3EAA45B950FA658 CRC64;

OY 4 KRRRRRRSG 13
Db 13 KRRRRRRRHG 22

Query Match 26.4%; Score 38; DB 1; Length 27;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
RL1 ARATH STANDARD; PRT; 25 AA.
ID P35015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L41.
GN (RP141-A OR AT1G56045 OR T6H22.15) AND (RP141-B OR AT3G08520 OR
GN T6G24.5 OR T6G24.5) AND (RP141-C OR AT3G11120 OR F9F8.7) AND
GN (RP141-D OR AT3G56020 OR F27K19.200).
OS Arabidopsis thaliana (Mouse-ear cress).
OS Nicotiana glauca (Common tobacco).
OS Gossypium hirsutum (Upland cotton).
OS Pisum sativum (Garden pea).
OS Hordeum vulgare (Barley). and
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702, 4097, 3635, 3888, 4513, 4530;
RN [1]
RP SEQUENCE FROM N.A. (AT1G56045).
RC SPECIES-A. thaliana; STRAIN-CV. COLUMBIA;
RX MEDLINE-21016719; Pubmed-1130712;
RA Theologos A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Eguu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marillat A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu T., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).

RN [2]
RP SEQUENCE FROM N.A. (AT3G08520; AT3G11120 AND AT3G56020).
RC SPECIES-A. thaliana; STRAIN-CV. COLUMBIA;
RX MEDLINE-21016720; Pubmed-1130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delenay M., Boutry M., Grivell L.A., Mache R., Pulgomech P.,
RA De Simone V., Choisy N., Artiguenave F., Robert C., Brothier P.,
RA Wincker P., Catolico L., Weissendbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmback E., Drzonek H., Erfle H., Jordan R., Bangert S.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Stimonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Schaefer M., Schoen O., Bargues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwelder B., Duchemant D.,
RA Cooke R., Laude M., Berger-Liauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Uterback T., Fujii C.Y., Shea T.P.,
RA Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pal G., Millscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsuno M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shingo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-A. thaliana; STRAIN-CV. COLUMBIA;
RX Cheuk R., Chen H., Kim C.J., Koeseema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamuya A., Karlin-Neumann G.,
RA Kawai I., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Tortum M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES-N. tabacum; STRAIN-CV. SR1; TISSUE-leaf;
RX Zhou X.R.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES-G. hirsutum; STRAIN-CV. DELTAPINE 62;
RX MEDLINE-95062728; Pubmed-7972506;
RA Turley R.B., Ferguson D.L., Meredith W.R.;
RT "Isolation and characterization of a cDNA encoding ribosomal protein
RT L41 from cotton (*Gossypium hirsutum* L.).";
RL Plant Physiol. 105:1449-1450(1994).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES-P. sativum; STRAIN-CV. LITTLE MARVEL; TISSUE-Root tip;
RX Woo H.H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES-H. vulgare; STRAIN-CV. Boml; TISSUE-Endosperm;
RX Rasmussen S.R.;
RT "Barley L41 ribosomal protein from immature endosperm.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES-O. sativa; STRAIN-CV. Nipponbare;

RA Buehl C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
 RA Zismann V., Pal G., Bowman C.L., Fujii C.Y., VanAken S.E.,
 RA Bowman C.L., Craven B., Utterback T.R., Khalak L., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.,
 RT "Oryza sativa chromosome 3 BAC OSINBA09J19 genomic sequence."
 CC Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE LALE FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AC009894; AAF02845.1; -
 DR EMBL: AC074395; AAG50829.1; -
 DR EMBL: AC009991; AAF01511.1; -
 DR EMBL: AL163832; CAB87856.1; -
 DR EMBL: AY058053; AAL24161.1; -
 DR EMBL: U26255; AAG7297.1; -
 DR EMBL: X75423; CA53175.1; -
 DR EMBL: L47967; AAA79268.1; -
 DR EMBL: AJ001160; CAA04564.1; -
 DR EMBL: AC084320; AAK09215.1; -
 DR PIR: S38425; S38425.
 KW Ribosomal protein.
 SQ SEQUENCE 25 AA; 3428 MW; 9AD629DA293C039E CRC64;
 QY 2 ERRKRRRRS 12
 Db 14 KRRKRRRRS 24
 RESULT 4
 RL1_HUMAN STANDARD: PRT; 25 AA.
 AC P28751;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L41 (HG12).
 GN RPL41.
 OS Homo sapiens (Human).
 OS Mus musculus (Mouse).
 OS Rattus norvegicus (Rat).
 OS Felis silvestris catus (Cat), and
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606, 10090, 10116, 9685, 31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human; TISSUE=Ovary;
 RX MEDLINE=92412140; PubMed=136959;
 RA Klaiding J., von der Kammer H., Scheit K.H.;
 RT "Characterization by cDNA cloning of the mRNA of a highly basic human
 RT protein homologous to the yeast ribosomal protein YL41";
 RL Biochem. Biophys. Res. Commun. 187:901-906(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RX MEDLINE=97446005; PubMed=9299532;
 RT Lee J.-H., Kim J.-M., Kim M.-S., Lee Y.-T., Marshak D.R., Bae Y.-S.;
 RT "The highly basic ribosomal protein L41 interacts with the beta
 RT subunit of protein kinase CKII and stimulates phosphorylation of DNA
 RT topoisomerase Ialpha by CKII.";

RL Biochem. Biophys. Res. Commun. 238:462-467(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Go H., Miyado K., Hasuwa H., Taniguchi S.;
 RT "Characterization of human ribosomal protein L41 genomic structure";
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Human;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman N., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Cader C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R., Cordy N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehaslaino M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McIay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.L.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senta H.R., Showkhen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas A.M., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmink L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
 RA Chan Y.-L., Wool I.G.;
 RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=C57BL/6;
 RA Rocha D., Anderson E., Botcherby M., Jordan B.;
 RL Submitted (Apr-1997) to the EMBL/Genbank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC SPECIES=F.s.catus;
 RX MEDLINE=96183078; PubMed=8607819;
 RA Starkey C.R., Menon R.P., Pradhu S., Levy L.S.;
 RT "Primary sequence and evolutionary conservation of ribosomal protein
 RT genes from the domestic cat.";
 RL Biochem. Biophys. Res. Commun. 220:648-652(1996).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC SPECIES=F.rubripes;
 RX MEDLINE=99177347; PubMed=10077531;
 RA Gellner K., Brenner S.;
 RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu
 RT rubripes.";
 RL Genome Res. 9:251-258(1999).
 CC -1- FUNCTION: INTERACTS WITH THE BETA SUBUNIT OF PROTEIN KINASE CKII
 CC AND STIMULATES PHOSPHORYLATION OF DNA TOPOISOMERASE II ALPHA BY
 CC CKII.
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE LALE FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: 2129663; CAA78306.1; -

DR EMBL: AF026844; AAB82715.1; -

DR EMBL: AB010874; BAA31508.1; -

DR EMBL: AL035562; CAB46994.1; -

DR EMBL: X82550; CAA57899.1; -

DR EMBL: U93862; AAB52254.1; -

DR EMBL: U22229; AAB01667.1; -

DR EMBL: AF056116; AAC34393.1; -

DR PIR: J01617; J01617.

DR PIR: S31691; S31691.

KW Ribosomal protein.

SQ SEQUENCE 25 AA; 3456 MW; 9AD629C69ECC039E CRC64;

Query Match 24.3%; Score 35; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKRRORRRS 12
Db 14 KRRKRRKRRS 24

RESULT 5
RL41.SOYBN STANDARD; PRT; 25 AA.

AC 049224;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L41.
GN RPL41.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosidia I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, ESSEX; TISSUE-Root;
RA Woo H.-H.;
RL Submitted (FEB-1998) TO THE EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE LA1E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: AF047051; AAC03557.1; -

DR Ribosomal protein.

KW RIBOSOMAL PROTEIN.

SQ SEQUENCE 25 AA; 3130 MW; AD933FB8E49ED6E CRC64;

Query Match 24.3%; Score 35; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 84;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERKRRORRRS 12
Db 14 KRRKRRKRRS 24

RESULT 5
PRTB.ACIGU STANDARD; PRT; 27 AA.
AC P02323; >

DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Protamine B (Sturine B) (Stellin B).
OS Acipenser guldenstadti (Caspian sturgeon) (Russian sturgeon), and
OS Acipenser stellatus (Sevruga).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenserinae; Acipenser.
OX NCBI_TaxID=7902, 7903;
RN [1]
RP SEQUENCE.
RC SPECIES-A.guldenstadti;
RA Yulikova E.P., Eysenko L.R., Baratova L.A., Belyanova L.P.,
RA Rybin V.K., Silaev A.B.;
RT "The primary structure of sturine B, a protamine from Caspian
RT sturgeon".
RL Bioorg. Khim. 2:1613-1617(1976).
RN [2]
RP SEQUENCE.
RC SPECIES-A.stellatus;
RA Rybin V.K., Yulikova E.P.;
RL Khim. Prirod. Soedin. 5:700-704(1979).
CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: TESTIS.
DR PIR: A02665; SRAPC.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Testis; DNA condensation; Nuclear protein.
SQ SEQUENCE 27 AA; 3707 MW; E300B46B173EC80 CRC64;

Query Match 23.6%; Score 34; DB 1; Length 27;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 4 KRRQRRRSG 13
Db 11 QRRRRRRHG 20

RESULT 7
RL41.SCHPO STANDARD; PRT; 25 AA.

AC 099710;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60S ribosomal protein L41.
GN (RPL41A OR SPAC36.13C) AND (RPL41B OR SPAC31F10.18C OR SPAC8F11.01C).
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A. (RPL41A).
RC STRAIN-972;
RA Gentles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (RPL41B).
RC STRAIN-972;
RA Botne G., Pohl T., McDougall R., Rajandream M.A., Barrell B.G.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L41 IN S.POMBE.
CC -1- SIMILARITY: BELONGS TO THE LA1E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; Z99167; CAB40152.1; -
 DR EMBL; Z69369; CAB40187.1; -
 DR EMBL; AL109738; CAB52162.1; -
 KM Ribosomal protein; Multigene family.
 SO SEQUENCE 25 AA; 3411 MW; 9EC629DD9ECC0535 CRC64;

Query Match 22.9%; Score 33; DB 1; Length 25;
 Best Local Similarity 54.5%; Pred. No. 1.5e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EKKRRRORRS 12
 :||:|:|:|
 DB 14 KRRKRRRARS 24

RESULT 8
 RL41_AGABI STANDARD; PRT; 25 AA.
 AC P78569;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE 60S ribosomal protein L41.
 GN R41.
 OS Agaricus bisporus (Common mushroom).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Agaricaceae; Agaricus.
 OX NCBI_TaxID=5341;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HORST U1;
 RX Sonnenberg A.S.M., de Groot P.W.J., Schaap P.J., Baars J.J.P.,
 RA Vissers J., van Griensven L.J.L.D.;
 RT "Isolation of expressed sequence tags of Agaricus bisporus and their
 RT assignment to chromosomes."
 RL Appl. Environ. Microbiol. 62:4542-4547(1996).
 CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X94764; CAA64390.1; -
 KM Ribosomal protein.
 SO SEQUENCE 25 AA; 3413 MW; A97629D0143C04EA CRC64;

Query Match 22.2%; Score 32; DB 1; Length 25;
 Best Local Similarity 60.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KRRRORRRS 12
 ||:|:|:|
 DB 15 KRRKRRRARS 24

RESULT 9
 RL41_QUESU STANDARD; PRT; 25 AA.
 AC O82713;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 60S ribosomal protein L41.
 GN RPL41.
 OS Quercus suber (Cork oak).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fagales; Fagaceae; Quercus.
 OX NCBI_TaxID=58331;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Huguet G., Pla M., Verdaguer D., Molinas M.;
 RT "Ribosomal proteins in Quercus suber";
 RT Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AJ001347; CAA04691.1; -
 KM Ribosomal protein.
 SO SEQUENCE 25 AA; 3337 MW; 6BC629DD9EDB3383 CRC64;

Query Match 22.2%; Score 32; DB 1; Length 25;
 Best Local Similarity 54.5%; Pred. No. 2.1e+02;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EKKRRRORRS 12
 :||:|:|:|
 DB 14 KRRKRRRARS 24

RESULT 10
 RL41_YEAST STANDARD; PRT; 25 AA.
 AC P05746;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 60S ribosomal protein L41 (YL41) (YL47).
 GN (RPL41 OR RPL47A OR YL41A OR YDL184C OR D1290) AND
 GN (RPL41B OR RPL47B OR YL41B OR YDL133BC).
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A. (RPL41A AND RPL41B).
 RX MEDLINE=90254826; PubMed=2187623;
 RA Suzuki K., Hashimoto T., Otake E.;
 RT "Yeast ribosomal proteins: XI. Molecular analysis of two genes
 RT encoding YL41, an extremely small and basic ribosomal protein, from
 RT Saccharomyces cerevisiae."
 RL Curr. Genet. 17:185-190(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (RPL41A).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=96021607; PubMed=8533471;
 RA Verhaesselt P., Voet M., Volckaert G.;
 RT "New open reading frames, one of which is similar to the n1v gene of
 RT Asotobacter vinelandii, found on a 12.5 kbp fragment of chromosome IV
 RT of Saccharomyces cerevisiae."
 RL Yeast 11:961-966(1995).
 RN [3]
 RP SEQUENCE FROM N.A. (RPL41B).
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97127826; PubMed=8972577;
 RA Woelfl S., Haneman V., Saluz H.P.;

```

RT "Analysis of a 26,756 bp segment from the left arm of yeast
RT chromosome IV."
RL Yeast 12:1549-1554(1996).
RP [4]
RA SEQUENCE.
RA Otake E., Migo K.-I., Itoh T.;
RT "Yeast ribosomal proteins: VIII. Isolation of two proteins and
RT sequence characterization of twenty-four proteins from cytoplasmic
RT ribosomes."
RL Mol. Gen. Genet. 195:544-546(1984).
CC -1- MISCELLANEOUS: THERE ARE TWO GENES FOR L41 IN YEAST.
CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16065; CAA34201.1; -
DR EMBL: X16066; CAA34202.1; -
DR EMBL: X83276; CAA58262.1; -
DR EMBL: Z74232; CAA98759.1; -
DR EMBL: Z74181; CAA98705.1; -
DR EMBL: Z74182; CAA98706.1; -
DR EMBL: X96876; CAA65626.1; -
DR PIR: S22247; R6BYA4.
DR PIR: S22246; R6BYA4.
DR SGD: S0002343; RPL41A.
DR SGD: S0002293; RPL41B.
DR Ribosomal protein; Multigene family.
KW RIBOSOMAL PROTEIN; 3337 MW; BD2629DD9ED85381 CRC64;
SQ SEQUENCE 25 AA: 3337 MW; BD2629DD9ED85381 CRC64;

Query Match 21.5%; Score 31; DB 1; Length 25;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERKRRRRRS 12
Db 14 KRKRKRRARS 24

RESULT 11
ID RL41_METJA STANDARD: PRT; 22 AA.
AC P54025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L41E.
GN RPL41E OR MJ0242.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake O., Fitzgerald L.M., Claydon R.A., Cocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöckel A.,
RA Usterbeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Nguyen D.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).

```

```

CC -1- SIMILARITY: BELONGS TO THE L41E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67480; AAB98230.1; -
DR TIGR: MJ0242; -
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 22 AA: 3089 MW; F948D74DD4EC98CB CRC64;

Query Match 20.1%; Score 29; DB 1; Length 22;
Best Local Similarity 41.7%; Pred. No. 4.5e+02;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 3 KRKRKRRRRSGS 14
Db 11 KRKRKRRRRARS 22

RESULT 12
ID TL16_SPIOL STANDARD: PRT; 29 AA.
AC P81834;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thylakoid lumenal 16.5 kDa protein (P16.5) (Fragment).
OS Spiroplasma oleraceae (Spiroplach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=35562;
RN 11
RP SEQUENCE.
RC TISSUE=leaf;
RX MEDLINE=98175931; PubMed=9506969;
RA Kieselbach T., Hagman A., Andersson B., Schroeder W.P.;
RA "The thylakoid lumen of chloroplasts - Isolation and
RT characterization."
RL J. Biol. Chem. 273:6710-6716(1998).
CC -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
KW Chloroplast; Thylakoid.
FT NON_TER 29
SQ SEQUENCE 29 AA: 3464 MW; 58B785764E2623E3 CRC64;

Query Match 19.4%; Score 28; DB 1; Length 29;
Best Local Similarity 55.6%; Pred. No. 8.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERKRRRRR 10
Db 18 DKRRRLRRQ 26

RESULT 13
ID NEUT_CAVPO STANDARD: PRT; 13 AA.
AC P32560;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neutroensin (NT).
GN NTS.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathii; Cavilidae; Cavia.

```

```

OX NCB1_Taxid-10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Small intestine;
RX MEDLINE=86248085; PubMed3067775;
RA Shaw C., Thim L., Conlon J.M.;
RT *Ser[1]neutensin: Isolation from guinea pig intestine.*;
RL FEBS Lett. 202;187-192(1986).
CC -1- FUNCTION: Smooth muscle-contracting peptide.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE NEUTROGENSIN FAMILY.
DR PIR; A53608; A53608.
KW Vasocactive.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 13 AA; 1680 MW; 4C8314644C4115B3 CRC64;

Query Match 18.8%; Score 27; DB 1; Length 13;
Best Local Similarity 71.4%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKRR 7
   || ||
Db 3 YENKSR 9

RESULT 14
UC06_MAIZE STANDARD; PRT; 15 AA.
AC P80612;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 1131)
DE (fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCB1_Taxid-4577;
RN [1]
RP SEQUENCE.
RC TISSUE=coleoptile;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RC Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.8, ITS MW IS: 71.0 kDa.
DR Maize-2DPAGE: P80612; COLEOPTILE.
DR MaizeDB: 123930; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1390 MW; 7005E22830F23D61 CRC64;

Query Match 18.8%; Score 27; DB 1; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 10 RRSGETDA 18
   || || ||
Db 7 RRGSGPDA 15

RESULT 15
HM12_CAEEL STANDARD; PRT; 27 AA.
AC P17487;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)

```

```

DE Homeobox protein ceh-12 (Fragment).
GN CEH-12.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCB1_Taxid-6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90245646; PubMed1970877;
RA Schaller D., Wittmann C., Spicher A., Mueller F., Tobler H.;
RT Cloning and analysis of three new homeobox genes from the nematode
RT Caenorhabditis elegans.*;
RL Nucleic Acids Res. 18:2033-2036(1990).
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X17076; CAA34928.1; -.
CC PIR: S09504; S09504.
CC HSSP: P02836; 1ENH.
CC TRANSFAC: T02987; -.
CC InterPro: IPR001356; Homeobox.
CC Pfam: PF00046; homeobox; 1.
CC PROSITE: PS00027; HOMEBOX_1; PARTIAL.
CC PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT NON_TER 1 1
FT DNA_BIND <1 1 HOMEBOX.
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3182 MW; FDAFD5D16F3870A5 CRC64;

Query Match 18.8%; Score 27; DB 1; Length 27;
Best Local Similarity 27.3%; Pred. No. 1e+03;
Matches 6; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 3 RKKRRQRRSGCTATAPASS 24
   : || : || : ||
Db 6 QNRMKMKKCPSPPIQSTSSS 27

```

Search completed: July 15, 2002, 13:57:02
 Job time: 315 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:57:48 ; Search time 39.48 seconds
(without alignments)
127.073 Million cell updates/sec

Title: US-09-712-819A-10
Perfect score: 144
Sequence: 1 YERKKRRRRSGSDATAPASSLOAED 29

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 12865

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rv1rus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44	30.6	26	13	Q90Y47	Q90Y47 cyanopsitta
2	44	30.6	26	13	Q90YV3	Q90YV3 ara maracn
3	44	30.6	26	13	Q90YV2	Q90YV2 cyanopsitta
4	42	29.2	23	8	Q35223	Q35223 oenothera b
5	37	25.7	13	12	Q65331	Q65331 autographa
6	35	24.3	25	13	Q90YS7	Q90YS7 ictalurus p
7	33	22.9	25	5	Q962S2	Q962S2 spodoptera
8	33	22.9	27	13	Q57542	Q57542 lampetra pl
9	33	22.9	27	13	Q57543	Q57543 lampetra pl
10	32	22.2	29	6	Q9TSS7	Q9TSS7 fellis silve
11	31	21.5	20	12	Q85617	Q85617 reovirus (t
12	31	21.5	22	12	Q85664	Q85664 reovirus (t
13	31	21.5	23	15	Q85496	Q85496 bovine leuk
14	31	21.5	29	5	Q17073	Q17073 antherea p
15	30	20.8	23	5	Q94781	Q94781 trypanosoma
16	30	20.8	23	11	Q35829	Q35829 rattus norv

17	30	20.8	26	4	Q9UGK1	Q9UGK1 homo sapien
18	30	20.8	22	12	Q9ORS7	Q9ORS7 hepatitis c
19	29.5	20.5	27	13	Q919P3	Q919P3 ara militar
20	29	20.1	20	4	Q9UNE8	Q9UNE8 homo sapien
21	29	20.1	26	12	Q56497	Q56497 hepatitis c
22	29	20.1	28	13	Q91395	Q91395 gallus gall
23	29	20.1	29	15	Q91Q15	Q91Q15 human immun
24	28	19.4	15	13	Q9PRM3	Q9PRM3 gallus gall
25	28	19.4	21	12	Q9JG30	Q9JG30 tt virus. o
26	28	19.4	21	12	Q9JG26	Q9JG26 tt virus. o
27	28	19.4	21	12	Q9JG23	Q9JG23 tt virus. o
28	28	19.4	23	8	Q9T2S6	Q9T2S6 nicotiana s
29	28	19.4	23	8	Q957T0	Q957T0 pinus mugo.
30	28	19.4	23	8	Q957S8	Q957S8 pinus sylve
31	28	19.4	26	2	Q33971	Q33971 streptomyce
32	28	19.4	26	11	Q99M13	Q99M13 mus musculu
33	28	19.4	27	2	Q9FB01	Q9FB01 streptomyce
34	28	19.4	27	5	Q95X71	Q95X71 caenorhabdi
35	28	19.4	28	2	Q951R3	Q951R3 streptomyce
36	28	19.4	28	13	Q9PRQ3	Q9PRQ3 micriturus n1
37	28	19.4	29	15	P89816	P89816 human immun
38	28	19.4	29	15	P89821	P89821 human immun
39	27.5	19.1	27	11	Q9R1H4	Q9R1H4 rattus norv
40	27	18.8	17	6	Q9TRR8	Q9TRR8 sus sp. ins
41	27	18.8	17	11	Q64450	Q64450 mus musculu
42	27	18.8	22	8	Q957T4	Q957T4 abies alba
43	27	18.8	22	13	Q13047	Q13047 xenopus lae
44	27	18.8	23	12	Q65291	Q65291 human adeno
45	27	18.8	25	4	Q96PE8	Q96PE8 homo sapien

ALIGNMENTS

RESULT	ID	PRELIMINARY:	PRT:	26 AA.
Q90Y47	Q90Y47	Q90Y47	PRT:	26 AA.
AC	Q90Y47	Q90Y47		
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)			
DE	CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).			
GN	CHD1-W.			
OS	Cyanopsitta spixii x Ara macacana.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Archosauria; Aves; Neognathae; Psittaciformes; Psittacidae.			
OX	NCBI_TaxID=157078;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Miyaki C.Y., Faria P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;			
RT	"The last Spix's Macaw and an Illiger's Macaw produced a hybrid.";			
RL	Conserv. Genet. 2:53-55(2001).			
DR	EMBL: AF276769; AAK96044.1; -.			
KW	Helicase.			
FT	NON_TER	1	1	
FT	NON_TER	26	26	
SQ	SEQUENCE	26 AA;	2973 MW;	P179E3EB84AC1297 CRC64;
Query Match				
Best Local Similarity 30.6%; Score 44; DB 13; Length 26;				
Matches 9; Conservative 52.9%; Pred. No. 18;				
Mismatches 3; Indels 5; Gaps 0;				
QY	2 ERKKRRRRSGSDA 18			
DB	11: 11111111			
DB	7 ERHRSRRYSGSDSDS 23			
RESULT	2			
Q90YV3	PRELIMINARY:			
AC	Q90YV3:	PRT:	26 AA.	
DT	01-DEC-2001 (TREMBLrel. 19, Created)			

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).
GN CHD1-W.
OS Ara maracana.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Psittaciformes; Psittacidae; Ara.
OX NCBI_TaxID=157077;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1, AND 2;
RA Miyaki C.Y., Farla P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;
RT "The last Spix's Macaw and an Illiger's Macaw produced a hybrid.";
RL Conserv. Genet. 2:53-55(2001).
DR EMBL; AF276764; AAK96039.1; -;
KW EMBL; AF276766; AAK96041.1; -;
FT Helicase.
FT NON_TER 1 26
FT NON_TER 26
SQ SEQUENCE 26 AA; 2973 MW; F179E3EB84AC1297 CRC64;

Query Match 30.6%; Score 44; DB 13; Length 26;
Best Local Similarity 52.9%; Pred. No. 18;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 ERKKRRRRSGSGTDA 18
DB 7 ERHRSRSGSDSDS 23

RESULT 3
O90VV2 PRELIMINARY; PRT; 26 AA.
AC O90VV2:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CHROMO-HELICASE DNA BINDING PROTEIN-1W (FRAGMENT).
GN CHD1-W.
OS Cyanopsitta spixii.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Psittaciformes; Psittacidae;
OC Cyanopsitta.
OX NCBI_TaxID=51902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1, AND 2;
RA Miyaki C.Y., Farla P.J., Griffiths R., Araujo J.C.C., Barros Y.M.;
RT "The last Spix's Macaw and an Illiger's Macaw produced a hybrid.";
RL Conserv. Genet. 2:53-55(2001).
DR EMBL; AF276767; AAK96042.1; -;
KW EMBL; AF276768; AAK96043.1; -;
FT Helicase.
FT NON_TER 1 26
FT NON_TER 26
SQ SEQUENCE 26 AA; 2973 MW; F179E3EB84AC1297 CRC64;

Query Match 30.6%; Score 44; DB 13; Length 26;
Best Local Similarity 52.9%; Pred. No. 18;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 ERKKRRRRSGSGTDA 18
DB 7 ERHRSRSGSDSDS 23

RESULT 4
O35223 PRELIMINARY; PRT; 23 AA.
AC O35223:
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN S19 (FRAGMENT).
GN RPS19.
OS Oenothera bertiana (Bertero's evening primrose).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=3950;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94250844; PubMed=8193306;
RA Bock H., Brennicke A., Schuster W.;
RT "Rps3 and rpl16 genes do not overlap in Oenothera mitochondria: CTS
as a potential translation initiation codon in plant mitochondria?";
RL Plant Mol. Biol. 24:811-818(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92107682;
RA Schuster W., Brennicke A.;
RT "RNA editing makes mistakes in plant mitochondria: editing looses
sense in transcripts of a rps19 pseudogene and in creating stop codons
in cox1 and rps3 mRNAs of Oenothera.";
RL Nucleic Acids Res. 19:6923-6928(1991).
DR EMBL; X69140; CAA48893.1; -;
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 1
SQ SEQUENCE 23 AA; 2764 MW; 08B5AB73D448C5CF CRC64;

Query Match 29.2%; Score 42; DB 8; Length 23;
Best Local Similarity 53.3%; Pred. No. 30;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 YERKKRRRRSGSG 15
DB 3 FTRKKRRSRRTGSG 17

RESULT 5
O65331 PRELIMINARY; PRT; 13 AA.
AC O65331:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 08, Last annotation update)
DE BASIC PROTEIN (FRAGMENT).
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C6;
RX MEDLINE=90218045; PubMed=2109042;
RA Hill-Perkins M.S., Possee R.D.;
RT "A baculovirus expression vector derived from the basic protein
promoter of Autographa californica nuclear polyhedrosis virus.";
RL J. Gen. Virol. 71:971-976(1990).
DR EMBL; D00864; BAA00738.1; -;
FT NON_TER 13 13
FT NON_TER 13
SQ SEQUENCE 13 AA; 1626 MW; 6C75B58A749D1414 CRC64;

Query Match 25.7%; Score 37; DB 12; Length 13;
Best Local Similarity 70.0%; Pred. No. 85;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 7 RORRRSGSGT 16
DB 4 RRRRSSTGT 13

RESULT 6
Q90YS7 PRELIMINARY; PRT; 25 AA.
ID 090YS7
DT 01-DEC-2001 (TRMBLrel. 19, Created)
DT 01-DEC-2001 (TRMBLrel. 19, Last sequence update)
DE RIBOSOMAL PROTEIN L41.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Siluriformes; Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RA Patterson A.P., Karst A., Liu Z.J.;
RT "Translational machinery of channel catfish: II. Complementary DNA and
RT Expression of the Complete Set of 47 60S Ribosomal Proteins."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF401598; AAK95170.1; -
SQ SEQUENCE 25 AA; 3357 MW; 9AD629C69EDC029E CRC64;

Query Match 24.3%; Score 35; DB 13; Length 25;
Best Local Similarity 54.5%; Pred. No. 3.2e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 ERKRRRRRS 12
Db 14 KRRKRRRRRS 24
:|:|:|:|:|

RESULT 7
Q962S2 PRELIMINARY; PRT; 25 AA.
ID Q962S2
DT 01-DEC-2001 (TRMBLrel. 19, Created)
DT 01-DEC-2001 (TRMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRMBLrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN L41.
OS Spodoptera frugiperda (Fall armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7108;
RN [1]
RP SEQUENCE FROM N.A.
RA Landais I., Ogilastro M., Mita K., Nohata J., Lopez-Farber M.,
RA Dionot-Cerutti M., Fournier P., Devauchelle G.;
RT "Full-length ribosomal protein sequence from an EST library of
RT Spodoptera frugiperda cells (Sf9)".
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF400204; AAK92176.1; -
SQ SEQUENCE 25 AA; 3399 MW; 9AD629DD9EC039E CRC64;

Query Match 22.9%; Score 33; DB 5; Length 25;
Best Local Similarity 54.5%; Pred. No. 6.1e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ERKRRRRRS 12
Db 14 KRRKRRRRRS 24
:|:|:|:|:|

RESULT 8
O57542 PRELIMINARY; PRT; 27 AA.
ID O57542
AC O57542
DT 01-JUN-1998 (TRMBLrel. 06, Created)
DT 01-JUN-1998 (TRMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TRMBLrel. 19, Last annotation update)

DE HOMEOBOX PROTEIN LPHOXIB (FRAGMENT).
OS Lampetra planeri (Brook lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7750;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98358009; PubMed=9694633;
RA Sharman A.C., Holland P.W.;
RT "Estimation of Hox gene cluster number in lampreys."
RL Int. J. Dev. Biol. 42:617-620(1998).
DR EMBL: AF044798; AAC0302.1; -
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3098 MW; FE1EE5E971DDCIDA CRC64;

Query Match 22.9%; Score 33; DB 13; Length 27;
Best Local Similarity 46.7%; Pred. No. 6.6e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKRRRRRSGSGTD 17
Db 9 RAPRRRRRAAQRD 23
| | | | | | | | | |

RESULT 9
O57543 PRELIMINARY; PRT; 27 AA.
ID O57543
AC O57543
DT 01-JUN-1998 (TRMBLrel. 06, Created)
DT 01-JUN-1998 (TRMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TRMBLrel. 19, Last annotation update)
DE HOMEOBOX PROTEIN LPHOXIC (FRAGMENT).
OS Lampetra planeri (Brook lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Lampetra.
OX NCBI_TaxID=7750;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98358009; PubMed=9694633;
RA Sharman A.C., Holland P.W.;
RT "Estimation of Hox gene cluster number in lampreys."
RL Int. J. Dev. Biol. 42:617-620(1998).
DR EMBL: AF044799; AAC03003.1; -
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 3190 MW; B86EE73971D0DC06 CRC64;

Query Match 22.9%; Score 33; DB 13; Length 27;
Best Local Similarity 46.7%; Pred. No. 6.6e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 RKRRRRRSGSGTD 17
Db 9 RAPRRRRRAAQRD 23
| | | | | | | | | |

RESULT 10
Q9TSS7 PRELIMINARY; PRT; 29 AA.
ID Q9TSS7
AC Q9TSS7
DT 01-MAY-2000 (TRMBLrel. 13, Created)
DT 01-MAY-2000 (TRMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRMBLrel. 19, Last annotation update)
DE WILM'S TUMOR PROTEIN 1 (FRAGMENT).
GN WTL1.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=21015404; PubMed=11130975;
RA Brouillette J.A.; Vénica P.J.;
RT "Estimate of nucleotide diversity in dogs with a pool-and-sequence
method."
RL Mamm. Genome 11:1079-1086(2000).
DR EMBL; AF201740; AAF19825.1; -.
FT NON_TER 1 1
SQ SEQUENCE 29 AA; 3594 MW; FDC8E57A0BAE35DD CRC64;

Query Match 22.2%; Score 32; DB 6; Length 29;
Best Local Similarity 66.7%; Pred. No. 9.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 5 KRRORRRSG 13
DB 1 KRRORRRHG 9

RESULT 11
085617 PRELIMINARY; PRT; 20 AA.
AC 085617;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE REOVIRUS SEROTYPE 1 L3 (FRAGMENT).
OS Reovirus (type 1 / strain Lang).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10884;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83017877; PubMed=7123853;
RA Gallard R.K.; Li J.K.; Keene J.D.; Joklik W.K.;
RT "The sequences at the terminal of four genes of the three reovirus
serotypes."
RL Virology 121:320-326(1982).
DR EMBL; J02297; AAA47233.1; -.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2177 MW; FA861D930E1FD8BA CRC64;

Query Match 21.5%; Score 31; DB 12; Length 20;
Best Local Similarity 33.3%; Pred. No. 9.4e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 5 KRRORRRSGGTDAT 19
DB 6 RRTKGKSSGKGDST 20

RESULT 12
085664 PRELIMINARY; PRT; 22 AA.
AC 085664;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE MAJOR CORE PROTEIN LAMBDA-1 (FRAGMENT).
OS Reovirus (type 3 / strain Dearing).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10886;
RN [1]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=DEARING;
RX MEDLINE=82217029; PubMed=7086967;
RA Kozak M.;
RT "Sequences of ribosome binding sites from the large size class of
reovirus mRNA."
RL J. Virol. 42:467-473(1982).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DEARING;
RX MEDLINE=83017876; PubMed=6927854;
RA Antczak J.B.; Chmelo R.; Pickup D.J.; Joklik W.K.;
RT "Sequences at both termini of the 10 genes of reovirus serotype 3
(strain Dearing)."
RL Virology 121:307-319(1982).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=DEARING;
RX MEDLINE=83017877; PubMed=7123853;
RA Gallard R.K.; Li J.K.; Keene J.D.; Joklik W.K.;
RT "The sequences at the terminal of four genes of the three reovirus
serotypes."
RL Virology 121:320-326(1982).
DR EMBL; J02317; AAA47254.1; -.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2462 MW; D6706A861D930E1F CRC64;

Query Match 21.5%; Score 31; DB 12; Length 22;
Best Local Similarity 33.3%; Pred. No. 1e+03;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 5 KRRORRRSGGTDAT 19
DB 6 RRTKGKSSGKGDST 20

RESULT 13
085496 PRELIMINARY; PRT; 23 AA.
AC 085496;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Bovine leukemia virus (BLV).
OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11901;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198886; PubMed=3033284;
RA Rice N.R.; Simek S.L.; Dubois G.C.; Showalter S.D.; Gilden R.V.;
RA Stephens R.M.;
RT "Expression of the bovine leukemia virus x region in virus-infected
RT cells."
RL J. Virol. 61:1577-1585(1987).
DR EMBL; M16017; AAA87336.1; -.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2877 MW; F85F7CBED5440B08 CRC64;

Query Match 21.5%; Score 31; DB 15; Length 23;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 KRRORRR 11
DB 3 KRRSRRR 10

RESULT 14
017073 PRELIMINARY; PRT; 29 AA.
AC 017073;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SILKNOTH (A. POLYPHEMUS) CHORION PROTEIN (FRAGMENT).
OS Anthraxa polyphemus (Polypheumus moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Saturniidae; Saturninae; Antherea;
 OX NCBI_TaxID=7120;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80090072; PubMed=519771;
 RA Jones C.W., Rosenthal N., Rodakis G.C., Kafatos F.C.;
 RT "Evolution of two major chorion multigene families as inferred from
 RT cloned cDNA and protein sequences.";
 RL Cell 18:1317-1332(1979).
 DR EMBL; K01842; AAA2785.1; -.
 KM Chorion.
 FT NON_TER 1 1
 FT NON_TER 29 29
 SQ SEQUENCE 29 AA; 2665 MW; 26483F783526E820 CRC64;

Query Match 21.5%; Score 31; DB 5; Length 29;
 Best Local Similarity 41.2%; Pred. 1.4e+03;
 Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 OY 11 RSGSGTDATAPASSLOA 27
 | | | : | | |
 Db 5 RLGGCGGISPAAELAA 21

RESULT 15
 Q94781
 ID Q94781 PRELIMINARY; PRT; 23 AA.
 AC Q94781;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
 DE HISTONE H2A (FRAGMENT).
 OS Trypanosoma cruzi.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Y;
 RA Tanaka T., Tanaka M.;
 RT "The parasite genome project: Trypanosoma cruzi.";
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; D87227; BAA13318.1; -.
 FT NON_TER 1 1
 FT NON_TER 23 AA; 2790 MW; 12E9ED7592E52045 CRC64;
 SQ SEQUENCE 23 AA; 2790 MW; 12E9ED7592E52045 CRC64;

Query Match 20.8%; Score 30; DB 5; Length 23;
 Best Local Similarity 66.7%; Pred. No. 1.5e+03;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 OY 3 RKKRRQRRR 11
 | | | | |
 Db 10 RDKRRRRR 18

Search completed: July 15, 2002, 13:57:49
 Job time: 337 sec

•

.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: July 15, 2002, 13:51:41 ; Search time 48.92 Seconds

(without alignments)
65.845 Million cell updates/sec

Title: US-09-712-819A-10

Perfect score: 144

Sequence: 1 YEKRRRRRRSSGTDATAPASSLQAE D 29

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues 272466

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 29

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	39.6	25	22	AA867911
2	56	38.9	18	17	AA86601
3	55	38.2	15	17	AA849577
4	55	38.2	15	18	AAW31207
5	55	38.2	15	18	AAW32916
6	55	38.2	15	18	AAW26448
7	55	38.2	15	19	AAW76153
8	55	38.2	16	18	AAW19818
9	55	38.2	22	22	AAE12486
10	55	38.2	22	22	AAE12487
11	55	38.2	24	15	AA849575

12	55	38.2	24	18	AAW31205	HIV-1 based transp
13	55	38.2	24	18	AAW32914	HIV-1 tat protein
14	55	38.2	24	18	AAW26446	HIV tat protein fr
15	55	38.2	24	19	AAW76151	HIV Type I TAT pro
16	55	38.2	26	15	AA849558	Residues 47-72 HIV
17	55	38.2	26	18	AAW31211	HIV-TAT protein tr
18	55	38.2	26	18	AAW26451	Peptide of the inv
19	55	38.2	26	19	AAW76157	HIV Type I TAT pro
20	55	38.2	27	20	AAW00349	Fragment of human
21	54	37.5	25	15	AA849567	Residues 38-62 HIV
22	54	37.5	25	18	AAW31216	HIV-TAT protein tr
23	54	37.5	25	18	AAW26456	Peptide of the inv
24	54	37.5	25	19	AAW76162	HIV Type I TAT pro
25	54	37.5	26	11	AAW03434	AA8 37-62 of HIV t
26	54	37.5	26	11	AAW02338	Tat-dm37-62, 46A1a
27	54	37.5	26	11	AAW02339	Tat-dm37-62, 41A1a
28	54	37.5	26	11	AAW04719	Tat-dm37-62, 40A1a
29	54	37.5	26	11	AAW04720	Tat-dm37-62, 39A1a
30	54	37.5	26	11	AAW02240	Tat-dm37-62, 38A1a
31	54	37.5	26	11	AAW02241	Tat-dm37-62, 37A1a
32	54	37.5	26	13	AAW30086	pH-dependent membr
33	54	37.5	26	21	AAW96570	HIV transactivator
34	54	37.5	29	19	AAW56399	Preferred signal s
35	54	37.5	29	21	AAW55820	Signal sequence fo
36	54	37.5	29	22	AAE11950	Membrane transport
37	53	36.8	13	21	AAW69771	RNA binding peptid
38	53	36.8	13	22	AAW06108	Peptide associated
39	53	36.8	21	22	AAW55701	Tat-CLASPcarboxyl
40	53	36.8	21	22	AAW57548	Human p75NTR modu
41	53	36.8	21	22	AAW57958	Beta-catenin depen
42	53	36.8	22	20	AAW33518	HIV Tat protein fr
43	52.5	36.5	22	21	AAW03933	Peptide associated
44	52.5	36.5	24	13	AAW27034	
45	52	36.1	21	22	AAW55700	

ALIGNMENTS

RESULT 1
AAB67911 standard; peptide; 25 AA.
XX
AC AAB67911:
XX
DT 29-JUN-2001 (first entry)
XX
DE HIV-1 tat derived peptide which is rich in basic amino acids.
XX
KW Heparin affinity regulatory peptide; HARP; angiogenic peptide; HIV;
KW Immune response; cellular growth; cellular regeneration; muscle cell;
KW Cytarisation; immunosuppressive disease; human immunodeficiency virus;
KW HIV infection; antiviral.
XX
OS Human immunodeficiency virus type 1.
XX
PN WO200127136-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000MO-FR02786.
XX
PR 12-OCT-1999; 99PR-0012714.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Barilault D, Achour A, Courty J;
XX
DR WPI; 2001-281970/29.
XX
PT New Heparin Affinity Regulatory Peptide angiogenic peptides, useful for
PT regenerating muscle cells, aiding cytarisation, detecting and
PT treating immunosuppressive diseases, stimulate immune response and

PT tissue regeneration -
XX
PS Disclosure: Page 3; 43pp; French.

CC AAB67904-16 represent peptides which rich in basic amino acids. The
CC specification describes heparin affinity regulatory peptide (HARP)
CC angiogenic peptides that stimulate immune response, cellular growth
CC and regeneration. The peptides are useful for regenerating cells, such
CC as muscle cells, and in aiding cicatrization, as well as in the
CC treatment of immunosuppressive diseases. The peptides amplify the
CC replication of human immunodeficiency virus (HIV) in vitro, and are
CC useful in the detection of HIV infections. When used with anti-viral
CC agents, they render the HIV more accessible to the antiviral agent.
CC and therefore are more easily destroyed.

XX
SQ Sequence 25 AA;

Query Match 39.6%; Score 57; DB 22; Length 25;
Best Local Similarity 66.7%; Pred. No. 0.4;
Matches 12; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 YERKKRRRRSGSGTDA 18
I | | | | | | | | | |
Db 7 YGRKKRRGRRPPGNGA 24

RESULT 2

AAR86601
ID AAR86601 standard: peptide; 18 AA.

XX
AC AAR86601;

XX 28-JUN-1996 (first entry)

XX HIV TAT cellular uptake region immunogen #1.

XX Immunogen; cellular uptake region; transactivating protein; TAT; HIV;

KW viremia; antibody.

XX
OS Human immunodeficiency virus type 1.

XX
PM WO9531999-A1.

XX
PD 30-NOV-1995.

XX
PF 16-MAY-1995; 95WO-US06077.

XX
PR 23-MAY-1994; 94US-0247991.

XX
PA (IMMU-) IMMUNOLOGY RES INST INC.

PI Culler MD, Goldstein G, Shenbagamurthi P;

DR WPI: 1996-049298/05.

XX
PT Immunogen containing cellular uptake region of viral TAT protein -

PT induces high antibody titre against TAT, partic. for preventing or

PT treating HIV infection

PS Claim 6; Page 45; 54pp; English.

XX
CC The sequences given in AAR86601-19 are immunogens derived from the
CC cellular uptake region of the transactivating protein, TAT of HIV.
CC These peptides are based on the region comprising amino acids 46-63
CC (see AAR86625) of the HIV TAT protein. These peptides are used to
CC protect against infection by HIV. They may also be used to reduce
CC viremia in already infected patients. These immunogens may be used
CC to identify, or generate, antibodies specific for TAT.

XX
SQ Sequence 18 AA;

Query Match 38.9%; Score 56; DB 17; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.39;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YERKKRRRRSGSGT 16
I | | | | | | | | | |
Db 2 YGRKKRRGRRPPGGS 17

RESULT 3

AAR49577
ID AAR49577 standard: peptide; 15 AA.

XX
AC AAR49577;

XX 23-SEP-1994 (first entry)

XX Residues tat47-58GGC HIV tat protein.

XX HIV; tat; transport; moiety; conjugate; cargo molecule;

KW cytoplasmic delivery; nuclear delivery; cysteine-rich region;

KW transactivation; disulphide aggregation.

XX
OS Human immunodeficiency virus type 1.

XX
PN WO9404686-A.

XX
PD 03-MAR-1994.

XX
PE 19-AUG-1993; 93WO-US07833.

XX
PR 21-AUG-1992; 92US-0934375.

XX
PA (BIOJ) BIOGEN INC.

PI Barsoum JG, Fawell SE, Pepinsky RB;

XX WPI: 1994-083202/10.

XX
PT Transport of cargo moieties into cells - using genetic fusions of

PT chemical conjugates comprising a portion of HIV tat protein as

PT transport peptide

PS Disclosure: Page 87; 153pp; English.

XX
CC The sequences given in AAR49574-77 are fragments of the HIV tat protein
CC which were used as transport moieties in the conjugate of the
CC invention. These conjugates are used as cargo molecules for the
CC efficient cytoplasmic and nuclear delivery of biologically active
CC non-tat proteins, nucleic acids and other molecules which are not
CC inherently capable of entering a target cell. The tat basic region
CC amino acid sequence carries cargo molecules by covalent attachment.
CC The reduced size of the transport peptides minimises interference
CC with the biological activity of the cargo molecule. By virtue of
CC the absence of the cysteine-rich region (residues 22-36) of the tat
CC protein, problems of spurious transactivation and disulphide
CC aggregation are solved. The reduced transport peptide size also
CC enhances uptake efficiency.

XX
SQ Sequence 15 AA;

Query Match 38.2%; Score 55; DB 15; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.45;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRRRSG 13
I | | | | | | | | | |
Db 1 YGRKKRRGRRPPG 13

RESULT 4

AAM31207

```

ID AAW31207 standard; peptide; 15 AA.
XX
AC AAW31207;
XX
DT 20-MAR-1998 (first entry)
XX
DE HIV-1 based transport polypeptide Tat47-58GGC.
XX
KM Human immunodeficiency virus; HIV Type 1; Tat protein;
KM cargo molecules; intracellular delivery; fusion protein;
KM therapeutic; prophylactic; diagnostic; transport polypeptide.
XX
OS Synthetic.
OS Human immunodeficiency virus type 1.
XX
FH Key location/Qualifiers
FT Region 3..11
FT Misc-difference 13..15
FT /label= basic_region
FT /label= spacer_residues
FT
XX US5674980-A.
XX
XX 07-OCT-1997.
XX
XX 21-DEC-1989; 89US-0454450.
XX
XX 28-APR-1994; 94US-0235403.
XX 21-DEC-1989; 89US-0454450.
XX 02-JAN-1991; 91US-0636662.
XX 21-AUG-1992; 92US-0934375.
XX 19-AUG-1993; 93MO-US07833.
XX 24-NOV-1993; 93US-0158015.
XX 25-MAY-1995; 95US-0450098.
XX
XX (BARS/) BARSOM J G.
XX (FAME/) FAMELL S E.
XX (FRAN/) FRANKEL A.
XX (PABO/) PABO.
XX (PEPI/) PEPINSKY R B.
XX
XX Barsom JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX WPI; 1997-502388/46.
XX
XX Fusion proteins containing truncated HIV tat sequences - useful for
XX intracellular delivery of viral repressor proteins
XX
XX Example 1: Column 55; 77pp; English.
XX
XX This sequence represents the transport polypeptide Tat47-58GGC which can
XX be used in a novel method for the delivery of biologically active cargo
XX molecules into the cytoplasm and nuclei of eukaryotic cells. The tat
XX protein from immunodeficiency virus (e.g. HIV-1, HIV-2, SIV) is readily
XX taken up into cells when present extracellularly. This polypeptide
XX has been constructed from the HIV-1 tat protein (see AAW31202) by
XX removing the cysteine-rich region and the carboxy-terminal domain while
XX retaining the basic region. Cargo molecules could be covalently linked
XX to the protein to form a fusion protein without the existing problems of
XX spurious trans-activation and disulphide aggregation. The reduced size
XX of transport polypeptides also minimises interference with the
XX biological activity of the cargo molecule. This is applicable for
XX therapeutic, prophylactic or diagnostic intracellular delivery of small
XX molecules and macromolecules e.g. proteins, nucleic acids and
XX polysaccharides.
XX
XX Sequence 15 AA:
XX

```

```

QY 1 YEKRRRRRRSG 13
DB 1 ygfkrkrqrrrp9 13
XX
XX RESULT 5
XX ID AAW32916 standard; peptide; 15 AA.
XX
XX AAW32916;
XX
XX 19-JAN-1998 (first entry)
XX
XX HIV-1 tat protein residues 47-58 plus Gly-Gly-Cys linker.
XX
XX Chemical conjugate; transport polypeptide; tat protein;
XX nucleic acid; delivery; cytoplasm; cell nucleus; therapy;
XX prophylaxis; diagnosis; spurious transactivation; HIV-1;
XX disulphide aggregation.
XX
XX Human immunodeficiency virus type 1.
XX
XX US5670617-A.
XX
XX 23-SEP-1997.
XX
XX 21-DEC-1989; 89US-0454450.
XX
XX 28-APR-1994; 94US-0235403.
XX 21-DEC-1989; 89US-0454450.
XX 02-JAN-1991; 91US-0636662.
XX 21-AUG-1992; 92US-0934375.
XX 19-AUG-1993; 93MO-US07833.
XX 24-NOV-1993; 93US-0158015.
XX 25-MAY-1995; 95US-0450246.
XX
XX (BARS/) BARSOM J G.
XX (FAME/) FAMELL S E.
XX (FRAN/) FRANKEL A.
XX (PABO/) PABO.
XX (PEPI/) PEPINSKY R B.
XX
XX Barsom JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
XX WPI; 1997-479523/44.
XX
XX Conjugate for intracellular delivery - comprising transport moiety
XX having amino acids 49-57 of human immunodeficiency virus tat protein
XX and nucleic acid cargo moiety
XX
XX Disclosure; Columns 57-58; 77pp; English.
XX
XX A novel covalently linked chemical conjugate, comprises a transport
XX polypeptide moiety consisting of at least residues 49-57 of human
XX immunodeficiency virus (HIV) tat protein, but not residues 22-36
XX and 73-86, e.g. the present peptide, and a cargo moiety comprising
XX a single or double stranded nucleic acid. The conjugate can be used
XX to deliver cargo molecules into the cytoplasm and nuclei of cells
XX for therapeutic, prophylactic and diagnostic applications. The HIV
XX tat protein is readily taken up into cells and the cell nucleus.
XX The reduced size of the transport polypeptides minimises
XX interference with the biological activity of the cargo molecule.
XX In addition, by virtue of the absence of the cysteine rich region
XX of the HIV tat protein, the transport polypeptides solve the
XX potential problems of spurious transactivation and disulphide
XX aggregation.
XX
XX Sequence 15 AA:
XX

```

Query Match 38.2%; Score 55; DB 18; Length 15;
 Best Local Similarity 84.6%; Pred. No. 0.45;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 38.2%; Score 55; DB 18; Length 15;
 Best Local Similarity 84.6%; Pred. No. 0.45;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSG 13
 | | | | | | | | | |
 Db 1 ygrkkrrqrtrrp 13

RESULT 6

AAW6448
 ID AAW6448 standard; Protein; 15 AA.
 XX
 AC AAW6448;
 XX
 DT 16-DEC-1997 (first entry)
 XX
 DE HIV tat protein fragment (aa47-58GCG).
 XX
 KM HIV: tat protein; transport protein; cargo delivery.
 XX
 OS Chimeric human immunodeficiency virus type 1;
 OS Chimeric synthetic.
 XX
 PN US5652122-A.
 XX
 PD 29-JUL-1997.
 XX
 PF 21-DEC-1989; 89US-0454450.
 XX
 PR 28-APR-1994; 94US-0235403.
 PR 21-DEC-1989; 89US-0454450.
 PR 02-JAN-1991; 91US-0636662.
 PR 21-AUG-1992; 92US-0934373.
 PR 19-AUG-1993; 93WO-US07833.
 PR 24-NOV-1993; 93US-0158015.
 PR 25-MAY-1995; 95US-0450257.
 XX
 PA (BARS/) BARSOUM J G.
 PA (FAWE/) FAWELL S E.
 PA (FRAN/) FRANKEL A.
 PA (PABO/) PABO C.
 PA (PEPI/) PEPINSKY R B.
 PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 PI WPI: 1997-392943/36.
 DR
 XX
 PT New DNA constructs for transporting molecules to cells - encode a
 PT fusion protein comprising a modified HIV tat protein and a
 PT carboxy-terminal cargo moiety
 PS
 XX
 PS Disclosure: Column 57-58; 76pp; English.
 CC This polypeptide comprises amino acids 47-58 of the HIV tat
 CC protein (see also AAW6443), plus a Gly-Gly-Cys C-terminal extension.
 CC Novel transport proteins comprise modified HIV tat covalently
 CC attached to a cargo molecule (see AAW6436-42). The modified tat
 CC protein (see AAW6444-49) is characterized by the presence of the tat
 CC basic region (amino acids 47-59), and the absence of the Cys-rich
 CC region (amino acids 22-36) and the exon 2-encoded C-terminal region
 CC (amino acids 73-86). These modifications solve the potential
 CC problems of spurious trans-activation and disulphide aggregation,
 CC while the reduced size of the transport protein minimises
 CC interference with the biological activity of the cargo molecule.
 CC DNA molecules that encode the modified tat fusion proteins are
 CC claimed and can be used to deliver polypeptides or nucleic acids to
 CC the cytoplasm or nuclei of cells in vivo or in vitro.
 XX
 SO Sequence 15 AA;

Query/Match 38.2%; Score 55; DB 18; Length 15;
 Best Local Similarity 84.6%; Pred. No. 0.45;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRSG 13
 | | | | | | | | | |
 Db 1 ygrkkrrqrtrrp 13

RESULT 7

AAW6153
 ID AAW6153 standard; protein; 15 AA.
 XX
 AC AAW6153;
 XX
 DT 24-NOV-1998 (first entry)
 XX
 DE HIV Type I TAT protein fragment tat47-58GCG.
 XX
 KM TAT protein; cargo molecule; therapy; diagnosis; transport protein;
 KM fusion protein; human papillomavirus E2 repressor; target cell.
 XX
 OS Human immunodeficiency virus type 1.
 OS
 XX
 PN US5804604-A.
 XX
 PD 08-SEP-1998.
 XX
 PF 25-MAY-1995; 95US-0450236.
 XX
 PR 28-APR-1994; 94US-0235403.
 PR 21-DEC-1989; 89US-0454450.
 PR 02-JAN-1991; 91US-0636662.
 PR 19-AUG-1993; 93WO-US07833.
 PR 24-NOV-1993; 93US-0158015.
 PR 25-MAY-1995; 95US-0450236.
 XX
 PA (BIOJ) BIOGEN INC.
 PI Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 PI WPI: 1998-505702/43.
 DR
 XX
 PT HIV tat-derived transport fusion proteins - used to deliver
 PT biological active molecules e.g. peptide(s) or nucleic acids,
 PT specifically into cytoplasm or nuclei of cells
 PS
 XX
 PS Disclosure: Column 57-58; 83pp; English.
 CC This sequence represents a human immunodeficiency virus (HIV) Type I
 CC TAT protein fragment, tat47-58GCG which is used in a method for the
 CC delivery of biologically active cargo molecules into the cytoplasm and
 CC nuclei of cells, for therapeutic, prophylactic or diagnostic purposes.
 CC This is accomplished by the presence of a small, basic section of tat
 CC transport protein of HIV. This is used as it is this protein which is
 CC observed to cause human cells in culture to take up HIV. The method
 CC involves the use of a cargo moiety in combination with a transport
 CC moiety usually in the form of a fusion protein. The cargo moiety is a
 CC human papillomavirus E2 repressor that retains its biological activity
 CC after delivery into a target cell and where the transport moiety is
 CC one of following HIV tat protein fragments (a) aa 47-58, (b) aa 47-72,
 CC (c) 38-72, (d) aa 38-58, (e) aa 37-58, (f) aa 1-21 and 38-72,
 CC (g) aa 47-62 or aa 38-62. The proteins allow delivery of specific
 CC peptides into cells at high concentrations due to use of existing
 CC transporters. Previous methods of delivery include bombardment and
 CC transfecting, which only allow a fraction of the cell population to be
 CC infected and can additionally damage cells as they cause physical opening
 CC of the cell walls/membranes to allow entry.
 XX
 SO Sequence 15 AA;

Query/Match 38.2%; Score 55; DB 19; Length 15;
 Best Local Similarity 84.6%; Pred. No. 0.45;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 1 YERKKRRQRRSG 13

Db 1 ygrkkrrqrrrrpg 13

RESULT 8

AAW19818 standard; Peptide; 16 AA.

AAW19818;

26-JAN-1998 (first entry)

Universal transfer vector sequence of chimeric coat protein.

Adenovirus; vector; coat protein; gene therapy; gene transfer; human; cancer; autoimmune disease; heart disease; infection; universal transfer vector; pBSS 75-100 pGS(tat).

Synthetic.

MO9720051-A2.

05-JUN-1997.

27-NOV-1996; 96MO-US19150.

21-AUG-1996; 96US-0701124.

28-NOV-1995; 95US-0563368.

21-AUG-1996; 96US-0700846.

(GENV-) GENVEC INC.

Brough DE, Kovesdi I, Wickham TJ;

WPI; 1997-310606/28.

N-PSDB; AAT70272.

Claim 7; Page 59-60; 121pp; English.

This peptide can be used as a universal transfer vector (UTV) and/or spacer sequence in novel chimeric adenovirus coat proteins (CPS). The peptide is encoded by a gene fragment (see AAT70272) of transfer plasmid pBSS 75-100 pGS(tat). Claimed UTVs/spacers are given in AAW19810-11, AAW19813-25, AAW19827, AAW19831-32 and AAW19834-43). Claimed CPS, such as fibre proteins, differ from wild-type CPS by the introduction of the UTV and/or spacer at or near the C-terminus. This imparts on the chimeric CP the ability to bind to and enter cells by means of a novel cell surface binding site. Adenoviral vectors comprising the chimeric CP are able to enter cells more efficiently than vectors comprising wild-type CP, especially at lower m.o.i. They are especially useful for gene therapy of e.g. cancers, genetic disorders, pathogenic infections, heart disease or autoimmune diseases.

Sequence 16 AA:

Query Match 38.2%; Score 55; DB 18; Length 16;

Best Local Similarity 78.6%; Pred. No. 0.48; Mismatches 11; Conservative 1; Indels 2; Gaps 0;

1 YEKRRRQRRRSG 14

3 ygrkkrrqrrrrpg 16

RESULT 9

AAE12486 standard; peptide; 22 AA.

AAE12486;

03-JAN-2002 (first entry)

Membrane transiting antiviral peptide bTAT-9.

Membrane transiting peptide; virucide; antiviral; Herpes Simplex Virus; HSV; HIV; Human Immunodeficiency Virus; CMV; cytomegalovirus.

Unidentified.

Key Location/Qualifiers

Modified-site 1/note="Biotin-aminohexanoyl Tyr"

MO200157072-A2.

09-AUG-2001.

06-FEB-2001; 2001MO-US03813.

07-FEB-2000; 2000US-180823P.

22-FEB-2000; 2000US-184057P.

(WISC) WISCONSIN ALUMNI RES FOUND.

Brandt C, Bulmann H;

WPI; 2001-638840/73.

Peptides comprising membrane transiting peptides useful for treating or preventing a virus infection, e.g., human immunodeficiency virus, herpes simplex virus and cytomegalovirus -

Claim 7; Page 15; 43pp; English.

The invention relates to peptides comprising membrane transiting peptides with antiviral properties. The peptides are useful for treating or preventing a virus infection in a warm blooded animal, e.g., enveloped viruses such as human immunodeficiency virus (HIV), herpes simplex virus (HSV), cytomegalovirus (CMV) and non-enveloped CC virus. Preferably, the peptides are useful for treating or preventing infections from one or more HSVs. The antiviral peptides are used for treating viral infections of the skin or part of the oral or genital cavity. The present sequence is membrane transiting antiviral peptide.

Sequence 22 AA;

Query Match 38.2%; Score 55; DB 22; Length 22;

Best Local Similarity 84.6%; Pred. No. 0.67; Mismatches 11; Conservative 0; Indels 2; Gaps 0;

1 YEKRRRQRRRSG 13

1 ygrkkrrqrrrrpg 13

RESULT 10

AAE12487 standard; peptide; 22 AA.

AAE12487;

03-JAN-2002 (first entry)

Membrane transiting antiviral peptide bTAT-9x.

Membrane transiting peptide; virucide; antiviral; Herpes Simplex Virus; HSV; HIV; Human Immunodeficiency Virus; CMV; cytomegalovirus.

Unidentified.

XX	
FH	Key
FT	Modified-site
FT	Location
xx	1
xx	/note= "BiotIn-aminohexanoyl Tyr"
PN	
xx	WO200157072-A2.
PD	
xx	09-AUG-2001.
PF	
xx	06-FEB-2001; 2001WO-US03813.
PR	
xx	07-FEB-2000; 2000US-180823P.
PR	
xx	22-FEB-2000; 2000US-184057P.
PA	
xx	(WISC) WISCONSIN ALUMNI RES FOUND.
PI	
xx	Brandt C, Bultmann H;
DR	
xx	WPI; 2001-638840/73.
PT	
PT	Peptides comprising membrane transiting peptides useful for treating or preventing a virus infection, e.g., human immunodeficiency virus, herpes simplex virus and cytomegalovirus -
xx	
PS	
xx	Claim 7; Page 15; 43pp; English.
CC	
CC	The invention relates to peptides comprising membrane transiting
CC	peptides with antiviral properties. The peptides are useful for
CC	treating or preventing a virus infection in a warm blooded animal,
CC	e.g., enveloped viruses such as human immunodeficiency virus (HIV),
CC	herpes simplex virus (HSV), cytomegalovirus (CMV) and non-enveloped
CC	viruses. Preferably, the peptides are useful for treating or preventing
CC	infections from one or more HSVs. The antiviral peptides are used for
CC	treating viral infections of the skin or part of the oral or genital
CC	cavity. The present sequence is membrane transiting antiviral peptide.
XX	
Sequence	22 AA;

```

Query Match 38.2% Score 55; DB 22; Length 22;
Best Local Similarity 84.6% Pred. No. 0.67;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 YEKRRRQRRRG 13
      1 | | | | | | | |
      1 ygrkkrrrrrrp 13

Db

RESULT 11
AAR49575
AAR49575 standard; peptide: 24 AA.
XX
AC AAR49575;
XX
DT 23-SEP-1994 (first entry)
XX
DE Residues 37-58GCC HIV tat protein.
XX
KW HIV; tat; transport; moiety; conjugate; cargo molecule;
KW cytoplasmic delivery; nuclear delivery; cysteine-rich region;
KW transactivation; disulphide aggregation.
XX
OS Human immunodeficiency virus type 1.
XX
XX MO9404686-A.
XX
XX 03-MAR-1994.
XX
PF 19-AUG-1993; 93WO-US07833.
XX
PR 21-AUG-1992; 92US-0934375.
XX
XX (BIOJ ) BIOGEN INC.

```

XX Barcoum JG, Fawell SE, Pepinsky RB;
XX WPI; 1994-083202/10.
XX
XX Transport of cargo moieties into cells - using genetic fusions of
PT chemical conjugates comprising a portion of HIV tat protein as
PT transport peptide
XX
XX
PS Disclosure: Page 86; 153pp; English.
XX
XX The sequences given in AAR49574-77 are fragments of the HIV tat protein
CC which were used as transport moieties in the conjugate of the
CC invention. These conjugates are used as cargo molecules for the
CC efficient cytoplasmic and nuclear delivery of biologically active
CC non-tat proteins, nucleic acids and other molecules which are not
CC inherently capable of entering a target cell. The tat basic region
CC amino acid sequence carries cargo molecules by covalent attachment.
CC The reduced size of the transport peptides minimises interference
CC with the biological activity of the cargo molecule.. By virtue of
CC the absence of the cysteine-rich region (residues 22-36) of the tat
CC protein, problems of spurious transactivation and disulphide
CC aggregation are solved. The reduced transport peptide size also
CC enhances uptake efficiency.
XX
SQ Sequence 24 AA;

```

Query Match Similarity      38.2%   Score 55; DB 15; Length 24;
Best Local Similarity      84.6%   Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY      1 YEKRRRRRRRSG 13
        | | | | | | | |
Db      10 YAKRRRRRRPQ 22

```

RESULT	12	
AAW31205		
ID	AAW31205	standard; peptide; 24 AA.
XX		
XX	AAW31205;	
XX		
DT	20-MAR-1998	(first entry)
XX		
DE	HIV-1 based transport polypeptide Tat38-58GCC.	
XX		
KW	Human immunodeficiency virus; HIV Type 1; Tat protein;	
KW	carbohydrate; intracellular delivery; fusion protein;	
KW	therapeutic; prophylactic; diagnostic; transport polypeptide	
XX		
OS	Synthetic.	
OS	Human immunodeficiency virus type 1.	
XX		
Key	Location/Qualifiers	
FH	Region	12..20
FT		/label= basic_domain
FT	Misc-difference	22..24
FT		/note= "spacer residues"
XX		
PN	US5674980-A.	
XX		
PD	07-OCT-1997.	
XX		
PF	21-DEC-1989;	89US-0454450.
XX		
PR	28-APR-1994;	94US-0235403.
PR	21-DEC-1989;	89US-0454450.
PR	02-JAN-1991;	91US-0636662.
PR	21-AUG-1992;	92US-0934375.
PR	19-AUG-1993;	93WO-US07833.
PR	24-NOV-1993;	93US-0158015.
PR	25-MAY-1995;	95US-0450098.


```
XX (BARS/) BARSOM J G.
PA (FAME/) FAMEL S E.
PA (FAME/) FAMEL S E.
PA (FRAN/) FRANKEL A.
PA (PABO/) PABO.
PA (PEPI/) PEPI R B.
XX BARSOM JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
DR WPI: 1997-502388/46.
XX
XX
XX Fusion proteins containing truncated HIV tat sequences - useful for
PT intracellular delivery of viral repressor proteins
PS Example 1; Column 55; 77pp; English.
XX
XX This sequence represents the transport polypeptide Tat38-58GCG which can
CC be used in a novel method for the delivery of biologically active cargo
CC molecules into the cytoplasm and nuclei of eukaryotic cells. The tat
CC protein from immunodeficiency virus (e.g. HIV-1, HIV-2, SIV) is readily
CC taken up into cells when present extracellularly. This polypeptide
CC has been constructed from the HIV-1 tat protein (see AAW31202) by
CC removing the cysteine-rich region and the carboxy-terminal domain while
CC retaining the basic region. Cargo molecules could be covalently linked
CC to the protein to form a fusion protein without the existing problems of
CC spurious trans-activation and disulphide aggregation. The reduced size
CC of transport polypeptides also minimises interference with the
CC biological activity of the cargo molecule. This is applicable for
CC therapeutic, prophylactic or diagnostic intracellular delivery of small
CC molecules and macromolecules e.g. proteins, nucleic acids and
CC polysaccharides.
XX
XX Sequence 24 AA:
SQ

Query Match 38.2%; Score 55; DB 18; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRQRRSG 13
   1 | | | | | | | | |
Db 10 YGRKKRRQRRPG 22

RESULT 13
AAW32914
ID AAW32914 standard; peptide; 24 AA.
XX
AC AAW32914;
XX
DT 19-JAN-1998 (first entry)
XX
XX HIV-1 tat protein residues 38-58 plus Gly-Gly-Cys linker.
DE
XX
XX Chemical conjugate: transport polypeptide; tat protein;
KW nucleic acid; delivery; cytoplasm; cell nucleus; therapy;
KW prophylaxis; diagnosis; spurious transactivation; HIV-1;
KW disulphide aggregation.
XX
XX Human immunodeficiency virus type 1.
OS
XX US5670617-A.
PN
XX
XX 23-SEP-1997.
PD
XX
XX 21-DEC-1989; 89US-0454450.
PF
XX
XX 28-APR-1994; 94US-0235403.
PR
XX 21-DEC-1989; 89US-0454450.
PR 02-JAN-1991; 91US-0636662.
PR 21-AUG-1992; 92US-0934375.
PR 19-AUG-1993; 93WO-US07833.
PR 24-NOV-1993; 93US-0158015.
```

```
PR 25-MAY-1995; 95US-0450246.
XX
XX (BARS/) BARSOM J G.
PA (FAME/) FAMEL S E.
PA (FRAN/) FRANKEL A.
PA (PABO/) PABO C.
PA (PEPI/) PEPI R B.
XX BARSOM JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
DR WPI: 1997-479523/44.
XX
XX
XX Conjugate for intracellular delivery - comprising transport moiety
PT having amino acids 49-57 of human immunodeficiency virus tat protein
PT and nucleic acid cargo moiety
PS Disclosure; Columns 55-56; 77pp; English.
XX
XX A novel covalently linked chemical conjugate, comprises a transport
CC polypeptide moiety consisting of at least residues 49-57 of human
CC immunodeficiency virus (HIV) tat protein, but not residues 22-36
CC and 73-86, e.g. the present peptide, and a cargo moiety comprising
CC a single or double stranded nucleic acid. The conjugate can be used
CC to deliver cargo moieties into the cytoplasm and nuclei of cells
CC for therapeutic, prophylactic and diagnostic applications. The HIV
CC tat protein is readily taken up into cells and the cell nucleus.
CC The reduced size of the transport polypeptides minimises
CC interference with the biological activity of the cargo molecule.
CC In addition, by virtue of the absence of the cysteine rich region
CC of the HIV tat protein, the transport polypeptides solve the
CC potential problems of spurious transactivation and disulphide
CC aggregation.
XX
XX Sequence 24 AA:
SQ

Query Match 38.2%; Score 55; DB 18; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRQRRSG 13
   1 | | | | | | | | |
Db 10 YGRKKRRQRRPG 22

RESULT 14
AAW26446
ID AAW26446 standard; Protein; 24 AA.
XX
AC AAW26446;
XX
DT 16-DEC-1997 (first entry)
XX
XX HIV tat protein fragment (aa37-58GCG).
DE
XX
XX HIV; tat protein; transport protein; cargo delivery.
KW Chimeric human immunodeficiency virus type 1;
OS Chimeric synthetic.
OS
XX US5652122-A.
PN
XX
XX 29-JUL-1997.
PD
XX
XX 21-DEC-1989; 89US-0454450.
PF
XX
XX 28-APR-1994; 94US-0235403.
PR
XX 21-DEC-1989; 89US-0454450.
PR 02-JAN-1991; 91US-0636662.
PR 21-AUG-1992; 92US-0934375.
PR 19-AUG-1993; 93WO-US07833.
PR 24-NOV-1993; 93US-0158015.
PR 25-MAY-1995; 95US-0450257.
```

XX (BARS/) BARSOUM J G.
 PA (FAME/) FAMELL S E.
 PA (FRAN/) FRANKEL A.
 PA (PABO/) PABO C.
 PA (PEPI/) PEPIINSKY R B.
 XX Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 PI WPI: 1997-392943/36.
 XX
 DR
 XX
 PT New DNA constructs for transporting molecules to cells - encode a
 PT fusion protein comprising a modified HIV tat protein and a
 PT carboxy-terminal cargo moiety
 XX
 PS Disclosure: Column 55-56; 76pp; English.
 XX
 CC This polypeptide comprises amino acids 37-58 of the HIV tat
 CC protein (see also AAW26443), plus a Gly-Gly-Cys C-terminal extension.
 CC Novel transport proteins comprise modified HIV tat covalently
 CC attached to a cargo molecule (see AAW26436-42). The modified tat
 CC protein (see AAW26444-49) is characterised by the presence of the tat
 CC basic region (amino acids 47-59), and the absence of the Cys-rich
 CC region (amino acids 73-86). These modifications solve the potential
 CC problems of spurious trans-activation and disulphide aggregation,
 CC while the reduced size of the transport protein minimises
 CC interference with the biological activity of the cargo molecule.
 CC DNA molecules that encode the modified tat fusion proteins are
 CC claimed and can be used to deliver polypeptides or nucleic acids to
 CC the cytoplasm or nuclei of cells in vivo or in vitro.
 XX
 SQ Sequence 24 AA:

Query Match 38.2%; Score 55; DB 18; Length 24;
 Best Local Similarity 84.6%; Pred. No. 0.73;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YEKRRRQRRRSG 13
 I | | | | | | | | | |
 DB 10 YGKRRRRRRRSG 22

RESULT 15
 AAW76151
 ID AAW76151 standard; protein; 24 AA.
 XX
 AC AAW76151;
 XX
 DT 24-NOV-1998 (first entry)
 XX
 DE HIV Type I Tat protein fragment aa. 38-58GGC.
 XX
 KW Tat protein; cargo molecule; therapy; diagnosis; transport protein;
 KW fusion protein; human papillomavirus E2 repressor; target cell.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 PN US5804604-A.
 XX
 PD 08-SEP-1998.
 XX
 PF 25-MAY-1995; 95US-0450236.
 XX
 PR 28-APR-1994; 94US-0235403.
 PR 21-DEC-1989; 89US-0434450.
 PR 02-JAN-1991; 91US-0636662.
 PR 19-AUG-1993; 93WO-US07833.
 PR 24-NOV-1993; 93US-0158015.
 PR 25-MAY-1995; 95US-0450236.
 XX
 PA (BIOJ) BIOGEN INC.

XX Barsoum JG, Fawell SE, Frankel A, Pabo C, Pepinsky RB;
 PI WPI: 1998-505702/43.
 XX
 DR
 XX
 PT HIV tat-derived transport fusion proteins - used to deliver
 PT biological active molecules e.g. peptide(s) or nucleic acids,
 PT specifically into cytoplasm or nuclei of cells
 XX
 PS Disclosure: Column 55-56; 83pp; English.
 XX
 CC This sequence represents a human immunodeficiency virus (HIV) Type I
 CC Tat protein fragment, tat38-58GGC which is used in a method for the
 CC delivery of biologically active cargo molecules into the cytoplasm and
 CC nuclei of cells, for therapeutic, prophylactic or diagnostic purposes.
 CC This is accomplished by the presence of a small, basic section of tat
 CC transport protein of HIV. This is used as it is this protein which is
 CC observed to cause human cells in culture to take up HIV. The method
 CC involves the use of a cargo moiety in combination with a transport
 CC moiety usually in the form of a fusion protein. The cargo moiety is a
 CC human papillomavirus E2 repressor that retains its biological activity
 CC after delivery into a target cell and where the transport moiety is
 CC one of following HIV tat protein fragments (a) aa 47-58, (b) aa 47-72,
 CC (c) 38-72, (d) aa 38-58, (e) aa 37-58, (f) aa 1-21 and 38-72,
 CC (g) aa 47-62 or aa 38-62. The proteins allow delivery of specific
 CC peptides into cells at high concentrations due to use of existing
 CC transporters. Previous methods of delivery include bombardment and
 CC transforming, which only allow a fraction of the cell population to be
 CC infected and can additionally damage cells as they cause physical opening
 CC of the cell walls/membranes to allow entry.
 XX
 SQ Sequence 24 AA:

Query Match 38.2%; Score 55; DB 19; Length 24;
 Best Local Similarity 84.6%; Pred. No. 0.73;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YEKRRRQRRRSG 13
 I | | | | | | | | | |
 DB 10 YGKRRRRRRRSG 22

Search completed: July 15, 2002, 13:51:41
 Job time: 230 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:52:08 ; Search time 21.58 Seconds

(without alignments)
32.824 Million cell updates/sec

Title: US-09-712-819A-10

Perfect score: 144
Sequence: 1 YEKRRRRRRSGSGTDA7APASSLQAE D 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 142465

Minimum DB seq length: 0
Maximum DB seq length: 29

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PC7US.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	38.9	18	5	PCT-US95-06077-3
2	56	38.9	26	5	PCT-US95-06077-24
3	55	38.2	15	1	US-08-450-257-6
4	55	38.2	15	1	US-08-450-246-6
5	55	38.2	15	1	US-08-450-098-6
6	55	38.2	15	1	US-08-451-233-6
7	55	38.2	15	1	US-08-450-236-6
8	55	38.2	15	4	US-08-235-403-6
9	55	38.2	24	1	US-08-450-257-4
10	55	38.2	24	1	US-08-450-246-4
11	55	38.2	24	1	US-08-450-098-4
12	55	38.2	24	1	US-08-451-233-4
13	55	38.2	24	1	US-08-450-236-4
14	55	38.2	24	1	US-08-235-403-4
15	55	38.2	25	5	PCT-US91-02942-42
16	55	38.2	26	1	US-08-450-257-48
17	55	38.2	26	1	US-08-450-246-48
18	55	38.2	26	1	US-08-450-098-48
19	55	38.2	26	1	US-08-451-233-48
20	55	38.2	26	1	US-08-450-236-48
21	55	38.2	26	4	US-08-235-403-48
22	54	37.5	25	1	US-08-450-257-53
23	54	37.5	25	1	US-08-450-246-53
24	54	37.5	25	1	US-08-450-098-53
25	54	37.5	25	1	US-08-451-233-53
26	54	37.5	25	1	US-08-450-236-53
27	54	37.5	25	4	US-08-235-403-53

28	54	37.5	26	6	5169933-2	Patent No. 5169933
29	54	37.5	29	2	US-08-928-958-9	Sequence 9, Appl
30	54	37.5	29	2	US-09-072-429-9	Sequence 9, Appl
31	54	37.5	29	4	US-09-411-706-5	Sequence 5, Appl
32	53	36.8	28	4	US-09-041-886-47	Sequence 47, Appl
33	51	35.4	14	5	PCT-US95-06077-4	Sequence 4, Appl
34	51	35.4	14	5	PCT-US85-06077-10	Sequence 10, Appl
35	51	35.4	14	5	PCT-US95-06077-12	Sequence 12, Appl
36	51	35.4	25	5	PCT-US91-02942-26	Sequence 26, Appl
37	51	35.4	28	4	US-09-041-886-40	Sequence 40, Appl
38	50	34.7	11	2	US-08-706-741B-54	Sequence 54, Appl
39	50	34.7	11	2	US-08-924-695A-54	Sequence 4, Appl
40	50	34.7	11	4	US-09-208-966-2	Sequence 2, Appl
41	50	34.7	11	4	US-09-296-089-37	Sequence 37, Appl
42	50	34.7	12	1	US-08-450-257-47	Sequence 47, Appl
43	50	34.7	12	1	US-08-450-246-47	Sequence 47, Appl
44	50	34.7	12	1	US-08-450-098-47	Sequence 47, Appl
45	50	34.7	12	1	US-08-451-233-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
PCT-US95-06077-3
Sequence 3, Application PC/TUS9506077
GENERAL INFORMATION:
APPLICANT: Immunobiology Research, Institute Inc.
TITLE OF INVENTION: Vaccine interdictio of Extracellular
TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus
TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar
TITLE OF INVENTION: Interellular Transactivating Strategies
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: Spring House Corporate Cntr, PO Box 457
CITY: Spring House
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06077
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/247,991
FILING DATE: 23-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: IR144PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-06077-3

Query Match 38.9%; Score 56; DB 5; Length 18;
Best Local Similarity 68.8%; Pred. No. 0.058;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 YEKRRRRRRSGSGT 16

Db 2 YGRRKRRRRRAPGS 17

RESULT 2

PCT-US95-06077-24

Sequence 24, Application PC/TUS9506077

GENERAL INFORMATION:

APPLICANT: Immunobiology Research, Institute Inc.

TITLE OF INVENTION: Vaccine Interdiction of Extracellular

TITLE OF INVENTION: Transactivating Proteins of Human Immunodeficiency Virus

TITLE OF INVENTION: and Other Chronically Infecting Viruses Employing Similar

TITLE OF INVENTION: Intracellular Transactivating Strategies

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Howson and Howson

STREET: Spring House Corporate Cntr, PO Box 457

CITY: Spring House

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19477

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06077

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: B170 CIP 2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-450-257-6

Query Match

Best Local Similarity 38.2%; Score 55; DB 1; Length 15;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10020

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/450,257

FILING DATE: 25-MAY-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/235,403

FILING DATE: 28-APR-1994

APPLICATION NUMBER: US 07/934,375

FILING DATE: 21-AUG-1992

APPLICATION NUMBER: US 07/098,766

FILING DATE: 28-JUL-1993

APPLICATION NUMBER: PCT/US93/07833

FILING DATE: 19-AUG-1993

APPLICATION NUMBER: US 07/454,450

FILING DATE: 21-DEC-1989

APPLICATION NUMBER: US 07/636,662

FILING DATE: 02-JAN-1991

APPLICATION NUMBER: US 08/158,015

FILING DATE: 24-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Haley Jr., James F.

REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: B170 CIP 2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 596-9000

TELEFAX: (212) 596-9090

TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-450-257-6

Query Match

Best Local Similarity 38.2%; Score 55; DB 1; Length 15;

Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

Db 1 YGRRKRRRRRSG 13

Y 1 YGRRKRRRRRSG 13

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,246
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-450-246-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.066;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YEKRRRRRRSG 13
DB 1 YGRRRRRRRRPG 13

RESULT 5
US-08-450-098-6
Sequence 6, Application US/08450098
Patent No. 5674980
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOOM, James G.
APPLICANT: FAMELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,233
FILING DATE: 25-MAY-1995

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,098
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-450-098-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.066;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YEKRRRRRRSG 13
DB 1 YGRRRRRRRRPG 13

RESULT 6
US-08-451-233-6
Sequence 6, Application US/08451233
Patent No. 5747641
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOOM, James G.
APPLICANT: FAMELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,233
FILING DATE: 25-MAY-1995

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-451-233-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.066;
Matches 11: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRRRRSG 13
| | | | | | | | | | | | | | | | |
DB 1 YGRKKRRRRRRPG 13

RESULT 7
US-08-450-236-6
Sequence 6, Application US/08450236
Patent No. 5804604
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOW, James G.
APPLICANT: FAMELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,236
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/235,403
FILING DATE: 28-APR-1994

APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993
APPLICATION NUMBER: US 07/454,450
FILING DATE: 21-DEC-1989
APPLICATION NUMBER: US 07/636,662
FILING DATE: 02-JAN-1991
APPLICATION NUMBER: US 08/158,015
FILING DATE: 24-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B170 CIP 2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-450-236-6

Query Match 38.2%; Score 55; DB 1; Length 15;
Best Local Similarity 84.6%; Pred. No. 0.066;
Matches 11: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YERKKRRRRRSG 13
| | | | | | | | | | | | | | | | |
DB 1 YGRKKRRRRRRPG 13

RESULT 8
US-08-235-403-6
Sequence 6, Application US/08235403
Patent No. 6316003
GENERAL INFORMATION:
APPLICANT: FRANKEL, Alan
APPLICANT: PABO, Carl
APPLICANT: BARSOW, James G.
APPLICANT: FAMELL, Stephen E.
APPLICANT: PEPINSKY, R. B.
TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,403
FILING DATE: 21-AUG-1992
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/934,375
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/098,766
FILING DATE: 28-JUL-1993
APPLICATION NUMBER: PCT/US93/07833
FILING DATE: 19-AUG-1993


```

?      REGISTRATION NUMBER: 27 794
?      REFERENCE/DOCKET NUMBER: B170 CIP 2
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (212) 596-9000
?      TELEFAX: (212) 596-9090
?      TELEBX: 14-8867
?      INFORMATION FOR SEQ. ID NO.: 4:
?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 24 amino acids
?      TYPE: amino acid
?      TOPOLOGY: linear
?      MOLECULE TYPE: peptide
US-08-450-246-4

```

Query Match	38.28;	Score 55;	DB 1;	Length 24;
Best Local Similarity	84.68;	Pred. No. 0.11;		
Matches 11; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

RESULT 11
 US-08-450-098-4
 Sequence 4, Application US/08450098
 Patent No. 5674980
 GENERAL INFORMATION:
 APPLICANT: FRANKEL, Alan
 APPLICANT: PABO, Carl
 APPLICANT: BARSOUM, James G.
 APPLICANT: FANELL, Stephen E.
 APPLICANT: PEKINSKY, R. B.
 TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FISH & NEAVE
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10020
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/450,098
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/235,403
 FILING DATE: 28-APR-1994
 APPLICATION NUMBER: US 07/934,375
 FILING DATE: 21-AUG-1992
 APPLICATION NUMBER: US 07/098,766
 FILING DATE: 28-JUL-1993
 APPLICATION NUMBER: PCT/US93/07833
 FILING DATE: 19-AUG-1993
 APPLICATION NUMBER: US 07/454,450
 FILING DATE: 21-DEC-1989
 APPLICATION NUMBER: US 07/636,662
 FILING DATE: 02-JAN-1991
 APPLICATION NUMBER: US 08/158,015
 FILING DATE: 24-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., James F.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: B170 CIP 2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 596-9000

```

? TELEFAX: (212) 596-9090
?
? TELEX: 14-8367
?
? INFORMATION FOR SEQ ID NO:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 24 amino acids
?
? TYPE: amino acid
?
? TOPOLOGY: linear
?
? MOLECULE TYPE: peptide
US-08-450-098-4

```

```

Query Match      38.28;  Score 55;  DB 1;  Length 24;
Best Local Similarity 84.68;
Matches 11;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0.
QY      1 YEKRRRRRRRSG 13
      | | | | | | | |
Db      10 YGRRRRRRRRPG 22

```

RESULT 12
 US-08-451-233-4
 Sequence 4: Application US/08451233
 Patent No. 5747641
 GENERAL INFORMATION:
 APPLICANT: FRANKEL, Alan
 APPLICANT: PABO, Carl
 APPLICANT: BARSOUX, James G.
 APPLICANT: FANDELL, Stephen E.
 APPLICANT: PERINSKY, R. B.
 TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
 NUMBER OF SEQUENCES: 69
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FISH & NEAVE
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10020
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/451,233
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/235,403
 FILING DATE: 28-APR-1994
 APPLICATION NUMBER: US 07/934,375
 FILING DATE: 21-AUG-1992
 APPLICATION NUMBER: US 07/098,766
 FILING DATE: 28-JUL-1993
 APPLICATION NUMBER: PCT/US93/07833
 FILING DATE: 19-AUG-1993
 APPLICATION NUMBER: US 07/454,450
 FILING DATE: 21-DEC-1989
 APPLICATION NUMBER: US 07/636,662
 FILING DATE: 02-JAN-1991
 APPLICATION NUMBER: US 08/158,015
 FILING DATE: 24-NOV-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., James F.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: B170 CIP 2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 596-9000
 TELEFAX: (212) 596-9090
 TELEX: 14-8367
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-451-233-4

Query Match 38.2%; Score 55; DB 1; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.11;
Matches 11: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRRG 13
| | | | | | | | | | | | | | | | | |
Db 10 YGKRRRQRRRG 22

RESULT 13
US-08-450-236-4
; Sequence 4, Application US/08450236
; Patent No. 5804604
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,236
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,403
; FILING DATE: 28-APR-1994
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELE: 14-8367
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-450-236-4

Query Match 38.2%; Score 55; DB 1; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.11;
Matches 11: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YERKKRRQRRRG 13
| | | | | | | | | | | | | | | | | |
Db 10 YGKRRRQRRRG 22

RESULT 14
US-08-235-403-4
; Sequence 4, Application US/08235403
; Patent No. 6316003
; GENERAL INFORMATION:
; APPLICANT: FRANKEL, Alan
; APPLICANT: PABO, Carl
; APPLICANT: BARSOUM, James G.
; APPLICANT: FAWELL, Stephen E.
; APPLICANT: PEPINSKY, R. B.
; TITLE OF INVENTION: TAT-DERIVED TRANSPORT POLYPEPTIDES
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,403
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/934,375
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/098,766
; FILING DATE: 28-JUL-1993
; APPLICATION NUMBER: PCT/US93/07833
; FILING DATE: 19-AUG-1993
; APPLICATION NUMBER: US 07/454,450
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: US 07/636,662
; FILING DATE: 02-JAN-1991
; APPLICATION NUMBER: US 08/158,015
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B170 CIP 2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELE: 14-8367
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-235-403-4

Query Match 38.2%; Score 55; DB 4; Length 24;
Best Local Similarity 84.6%; Pred. No. 0.11;
Matches 11: Conservative 0; Mismatches 2; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:52:49 ; Search time 75.67 Seconds
(without alignments)
2.260 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FRLTIS 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents-AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/CTRUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	25	5	PCT-US91-02942-26 Sequence 26, Appl
2	32	100.0	25	5	PCT-US91-02942-42 Sequence 42, Appl
3	32	100.0	32	2	US-08-470-139-13 Sequence 13, Appl
4	32	100.0	32	4	US-08-569-147-40 Sequence 40, Appl
5	32	100.0	32	4	US-09-347-061-13 Sequence 13, Appl
6	32	100.0	32	4	US-09-425-638A-92 Sequence 92, Appl
7	32	100.0	32	4	US-09-425-638A-94 Sequence 94, Appl
8	32	100.0	32	4	US-09-543-004-92 Sequence 92, Appl
9	32	100.0	32	4	US-09-543-004-94 Sequence 94, Appl
10	32	100.0	50	5	PCT-US91-02942-8 Sequence 8, Appl
11	32	100.0	50	5	PCT-US91-02942-9 Sequence 9, Appl
12	32	100.0	64	2	US-08-765-179B-10 Sequence 10, Appl
13	32	100.0	64	2	US-08-765-179B-14 Sequence 14, Appl
14	32	100.0	70	3	US-08-554-840-9 Sequence 9, Appl
15	32	100.0	76	4	US-08-851-362D-21 Sequence 21, Appl
16	32	100.0	80	3	US-08-554-840-10 Sequence 10, Appl
17	32	100.0	80	3	US-08-554-840-12 Sequence 12, Appl
18	32	100.0	80	3	US-08-554-840-13 Sequence 13, Appl
19	32	100.0	80	3	US-08-554-840-15 Sequence 15, Appl
20	32	100.0	93	3	US-08-783-853A-35 Sequence 35, Appl
21	32	100.0	95	2	US-08-290-592E-19 Sequence 19, Appl
22	32	100.0	95	2	US-08-290-592E-33 Sequence 33, Appl
23	32	100.0	95	5	PCT-US95-10053-16 Sequence 16, Appl
24	32	100.0	95	5	PCT-US95-10053-30 Sequence 30, Appl
25	32	100.0	95	5	PCT-US96-0944B-19 Sequence 19, Appl
26	32	100.0	95	5	PCT-US96-0944B-33 Sequence 33, Appl
27	32	100.0	96	3	US-08-466-36B-6 Sequence 6, Appl

28	32	100.0	103	4	US-09-240-274-42 Sequence 42, Appl
29	32	100.0	104	1	US-08-276-852-106 Sequence 106, App
30	32	100.0	104	1	US-08-899-575-106 Sequence 106, App
31	32	100.0	104	1	US-08-899-575-106 Sequence 106, App
32	32	100.0	104	5	PCT-US95-08743-106 Sequence 89, Appl
33	32	100.0	105	1	US-08-276-852-89 Sequence 89, Appl
34	32	100.0	105	1	US-08-899-575-89 Sequence 89, Appl
35	32	100.0	105	1	US-08-899-575-89 Sequence 89, Appl
36	32	100.0	105	5	PCT-US95-08743-89 Sequence 89, Appl
37	32	100.0	106	1	US-07-634-278-7 Sequence 7, Appl
38	32	100.0	106	1	US-07-634-278-8 Sequence 8, Appl
39	32	100.0	106	1	US-08-477-728-7 Sequence 7, Appl
40	32	100.0	106	1	US-08-477-728-8 Sequence 8, Appl
41	32	100.0	106	1	US-08-276-852-83 Sequence 83, Appl
42	32	100.0	106	1	US-08-276-852-85 Sequence 85, Appl
43	32	100.0	106	1	US-08-474-040-7 Sequence 7, Appl
44	32	100.0	106	1	US-08-474-040-8 Sequence 8, Appl
45	32	100.0	106	1	US-08-487-200-7 Sequence 7, Appl

ALIGNMENTS

RESULT 1
PCT-US91-02942-26
Sequence 26, Application PC/TUS9102942
GENERAL INFORMATION:
APPLICANT: ROTHLEIN, ROBERT
APPLICANT: ADAIR, JOHN R
TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Ave. NW Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02942
FILING DATE: 19910429
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9009549.8
FILING DATE: 27-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAM L
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1011.0586600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-02942-26

Query Match 100.0%; Score 32; DB 5; Length 25;
Best local Similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FRLTIS 7
|||||||

Db 8 FTLTSS 14

RESULT 2

PCT-US91-02942-42

Sequence 42, Application PC/TUS9102942

GENERAL INFORMATION:

APPLICANT: ROTHLEIN, ROBERT

APPLICANT: ADAIR, JOHN R

APPLICANT: ATHWAL, DILJEET S

TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY

NUMBER OF SEQUENCES: 102

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sherre, Kessler, Goldstein & Fox

STREET: 1225 Connecticut Ave. NW Suite 300

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02942

FILING DATE: 19910429

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9009549, 8

FILING DATE: 27-APR-1990

ATTORNEY/AGENT INFORMATION:

NAME: FOX, SAM L

REGISTRATION NUMBER: 30,353

REFERENCE/DOCKET NUMBER: 1011.0586600

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 466-0800

TELEFAX: (202) 833-8716

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:

LENGTH: 25 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US91-02942-42

Query Match

Best Local Similarity 100.0%; Score 32; DB 5; Length 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7

Db 8 FTLTSS 14

RESULT 3

US-08-470-139-13

Sequence 13, Application US/08470139

Patent No. 5998586

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Interleukin-5 specific recombinant antibodies

NUMBER OF SEQUENCES: 28

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470.139

FILING DATE: 06 JUNE-1995

CLASSIFICATION: 536

Query Match

Best Local Similarity 100.0%; Score 32; DB 4; Length 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7

Db 15 FTLTSS 21

ATTORNEY/AGENT INFORMATION:

NAME: TRUJILLO, DOREEN YATKO

REGISTRATION NUMBER: 35,719

REFERENCE/DOCKET NUMBER: CARP-0044

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-470-139-13

Query Match

Best Local Similarity 100.0%; Score 32; DB 2; Length 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7

Db 15 FTLTSS 21

RESULT 4

US-08-569-147-40

Sequence 40, Application US/08569147

Patent No. 6180377

GENERAL INFORMATION:

TITLE OF INVENTION: HUMANISED ANTIBODIES

NUMBER OF SEQUENCES: 95

CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/569,147

FILING DATE: 25-March-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Trujillo, Doreen Yatko

REGISTRATION NUMBER: 35,719

REFERENCE/DOCKET NUMBER: CARP-0047

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 40:

SEQUENCE CHARACTERISTICS:

LENGTH: 32 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-569-147-40

```
RESULT 5
US-09-347-061-13
; Sequence 13, Application US/09347061
; Patent No. 6316227
; GENERAL INFORMATION:
; APPLICANT: Bodmer, Mark
; APPLICANT: Athwal, Diljeet Singh
; APPLICANT: Emtage, John Spencer
; TITLE OF INVENTION: Interleukin-5 Specific Recombinant Antibodies
; FILE REFERENCE: CARP-0071
; CURRENT APPLICATION NUMBER: US/09/347,061
; CURRENT FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Consensus
US-09-347-061-13
```

```
Query Match          100.0%; Score 32; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Indels 0; Gaps 0;
```

```
QY 1 FTLTSS 7
    |||||
Db 15 FTLTSS 21
```

```
RESULT 6
US-09-425-638A-92
; Sequence 92, Application US/09425638A
; Patent No. 6342587
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630
; CURRENT APPLICATION NUMBER: US/09/425,638A
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 92
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-425-638A-92
```

```
Query Match          100.0%; Score 32; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FTLTSS 7
    |||||
Db 15 FTLTSS 21
```

```
RESULT 7
US-09-425-638A-94
; Sequence 94, Application US/09425638A
; Patent No. 6342587
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630
; CURRENT APPLICATION NUMBER: US/09/425,638A
```

```
; CURRENT FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 94
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-425-638A-94
```

```
Query Match          100.0%; Score 32; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FTLTSS 7
    |||||
Db 15 FTLTSS 21
```

```
RESULT 8
US-09-543-004-92
; Sequence 92, Application US/09543004
; Patent No. 6346249
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630.1
; CURRENT APPLICATION NUMBER: US/09/543,004
; CURRENT FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: 09/425,638
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 92
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-543-004-92
```

```
Query Match          100.0%; Score 32; DB 4; Length 32;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 FTLTSS 7
    |||||
Db 15 FTLTSS 21
```

```
RESULT 9
US-09-543-004-94
; Sequence 94, Application US/09543004
; Patent No. 6346249
; GENERAL INFORMATION:
; APPLICANT: Carlos F. Barbas III, Christoph Rader, Gerd Ritter, Sydney Welt and
; APPLICANT: Lloyd J. Old
; TITLE OF INVENTION: A33 ANTIGEN SPECIFIC IMMUNOGLOBULIN PRODUCTS AND USES THEREOF
; FILE REFERENCE: LUD 5630.1
; CURRENT APPLICATION NUMBER: US/09/543,004
; CURRENT FILING DATE: 2000-04-04
; PRIOR APPLICATION NUMBER: 09/425,638
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 129
; SEQ ID NO 94
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-543-004-94
```

```
Query Match          100.0%; Score 32; DB 4; Length 32;
```

Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||||
DB 15 FTLTSS 21

RESULT 10
PCT-US91-02942-8

; Sequence 8, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: ATHMAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/02942
; FILING DATE: 19910429
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9009549.8
; FILING DATE: 27-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: FOX, SAM L
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 1011.0586600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US91-02942-8

Query Match 100.0%; Score 32; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||||
DB 26 FTLTSS 32

RESULT 11
PCT-US91-02942-9

; Sequence 9, Application PC/TUS9102942
; GENERAL INFORMATION:
; APPLICANT: ROTHLEIN, ROBERT
; APPLICANT: ADAIR, JOHN R
; APPLICANT: ATHMAL, DILJEET S
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02942
FILING DATE: 19910429
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9009549.8
FILING DATE: 27-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: FOX, SAM L
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 1011.0586600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US91-02942-9

Query Match 100.0%; Score 32; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||||
DB 26 FTLTSS 32

RESULT 12
US-08-765-179B-10
; Sequence 10, Application US/08765179B
; Patent No. 5854027
; GENERAL INFORMATION:
; APPLICANT: STEINBACHER, Stefan
; TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY
; TITLE OF INVENTION: OF ANTIBODIES
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,179B
; FILING DATE: 14-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/02626
; FILING DATE: 06-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 25 115.7
; FILING DATE: 15-JUL-1994

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-179B-10

Query Match 100.0%; Score 32; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||
Db 40 FTLTSS 46

RESULT 13
US-08-765-179B-14
Sequence 14, Application US/08765179B
Patent No. 5854027
GENERAL INFORMATION:
APPLICANT: STEIPE, Boris
APPLICANT: STEINBACHER, Stefan
TITLE OF INVENTION: PROCESS FOR MODIFYING THE STABILITY
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/765,179B
FILING DATE: 14-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02626
FILING DATE: 06-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 44 25 115.7
FILING DATE: 15-JUL-1994
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-765-179B-14

Query Match 100.0%; Score 32; DB 2; Length 64;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||
Db 40 FTLTSS 46

RESULT 14
US-08-554-840-9
Sequence 9, Application US/08554840

Patent No. 6001358
GENERAL INFORMATION:
APPLICANT: BLACK, Amelia
APPLICANT: HANNA, Nabil
APPLICANT: PADLAN, Eduardo A.
APPLICANT: NEWMAN, Roland A.
TITLE OF INVENTION: HUMANIZED ANTIBODIES TO HUMAN gp39,
COMPOSITIONS CONTAINING AND THERAPEUTIC USE THEREOF
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 2213-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,840
FILING DATE: 07-NOV-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 012712-127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 70 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-554-840-9

Query Match 100.0%; Score 32; DB 3; Length 70;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||
Db 53 FTLTSS 59

RESULT 15
US-08-851-362D-21
Sequence 21, Application US/08851362D
Patent No. 6235883
GENERAL INFORMATION:
APPLICANT: Jakobovits, Aya
APPLICANT: Yang, Xiao-Dong
APPLICANT: Gallo, Michael
APPLICANT: Jia, Xiao-Chi
TITLE OF INVENTION: Human Monoclonal Antibodies to Epidermal
TITLE OF INVENTION: Growth Factor Receptor
FILE REFERENCE: Cell 4.20
CURRENT APPLICATION NUMBER: US/08/851,362D
CURRENT FILING DATE: 1997-05-05
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21
LENGTH: 76
TYPE: PRT
ORGANISM: human
US-08-851-362D-21

Query Match 100.0%; Score 32; DB 4; Length 76;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 FTLTSS 7
 |||||
 Db 52 FTLTSS 58

Search completed: July 15, 2002, 12:59:16
 Job time: 387 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:53:14 ; Search time 95.45 seconds

(without alignments)
7.047 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTLTSS 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	54	2 JT0521	Ig kappa chain V-I
2	32	100.0	71	2 S21526	Ig kappa chain V-I
3	32	100.0	77	2 D30502	Ig kappa chain V-I
4	32	100.0	77	2 S70443	Ig kappa chain V-I
5	32	100.0	78	2 S34102	Ig kappa chain V-I
6	32	100.0	79	2 S24215	Ig kappa chain V-I
7	32	100.0	81	2 PH1048	Ig light chain V-I
8	32	100.0	83	2 S78489	Ig kappa chain V-I
9	32	100.0	84	2 S34099	Ig kappa chain V-I
10	32	100.0	86	2 S16826	Ig kappa chain V-I
11	32	100.0	86	2 S16834	Ig kappa chain V-I
12	32	100.0	86	2 S16836	Ig kappa chain V-I
13	32	100.0	86	2 S34086	Ig kappa chain V-I
14	32	100.0	86	2 S16840	Ig kappa chain V-I
15	32	100.0	86	2 S16837	Ig kappa chain V-I
16	32	100.0	86	2 S16833	Ig kappa chain V-I
17	32	100.0	86	2 S16830	Ig kappa chain V-I
18	32	100.0	86	2 S16824	Ig kappa chain V-I
19	32	100.0	86	2 S16829	Ig kappa chain V-I
20	32	100.0	87	2 S34098	Ig kappa chain V-I
21	32	100.0	87	2 S21523	Ig kappa chain V-I
22	32	100.0	87	2 S34097	Ig kappa chain V-I
23	32	100.0	87	2 S34084	Ig kappa chain V-I
24	32	100.0	87	2 S34083	Ig kappa chain V-I
25	32	100.0	88	2 S21528	Ig kappa chain V-I
26	32	100.0	88	2 S21525	Ig kappa chain V-I
27	32	100.0	88	2 S21520	Ig kappa chain V-I
28	32	100.0	88	2 S21522	Ig kappa chain V-I
29	32	100.0	88	2 S34104	Ig kappa chain V-I

30	32	100.0	91	2 PH1071	Ig light chain V-I
31	32	100.0	91	2 S37525	Ig kappa chain V-I
32	32	100.0	91	2 S37527	Ig kappa chain V-I
33	32	100.0	91	2 S37521	Ig kappa chain V-I
34	32	100.0	91	2 S37511	Ig kappa chain V-I
35	32	100.0	91	2 S37515	Ig kappa chain V-I
36	32	100.0	92	2 S37533	Ig kappa chain V-I
37	32	100.0	92	2 S37530	Ig kappa chain V-I
38	32	100.0	92	2 S37529	Ig kappa chain V-I
39	32	100.0	92	2 S37523	Ig kappa chain V-I
40	32	100.0	92	2 S37535	Ig kappa chain V-I
41	32	100.0	92	2 S37509	Ig kappa chain V-I
42	32	100.0	92	2 S37512	Ig kappa chain V-I
43	32	100.0	92	2 S37531	Ig kappa chain V-I
44	32	100.0	92	2 S37534	Ig kappa chain V-I
45	32	100.0	92	2 S37532	Ig kappa chain V-I

ALIGNMENTS

RESULT 1
JT0521
Ig kappa chain V-III region (CPI) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 23-Oct-1992 #sequence_revision 23-Oct-1992 #text_change 09-May-1997
C:Accession: JT0521
R:Anker, R.; Conley, M.E.; Pollock, B.A.
J. Exp. Med. 169, 2109-2119, 1989
A:Title: Clonal diversity in the B cell repertoire of patients with X-linked agammaglobulinemia
A:Reference number: JT0511; MUID:89279157
A:Accession: JT0521
A:Molecule type: mRNA
A:Residues: 1-54 <ANK>
A:Note: The sequence shown here is one of eight productive V-D-J mu chain rearrangements
C:Superfamily: Immunoglobulin V region, immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-49/Domain: V region <VRE>
F:50-54/Domain: J region <JRE>

Query Match 100.0%; Score 32; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
Db 23 FTLTSS 29

RESULT 2
S21526
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 25-Oct-1996 #text_change 23-Jul-1999
C:Accession: S34082; S21526
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34082
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <MA2>
A:Cross-references: EMBL:X66042; NID:g33318; PIDN:CAA46841.1; PID:g33319
C:Superfamily: immunoglobulin V region, immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 32; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 64 FTLTSS 70

RESULT 3

D30502
Ig kappa chain V region (D44) - mouse
C:Species: Mus musculus (house mouse)
C>Date: 03-Nov-1988 #sequence_revision 03-Aug-1992 #text_change 23-Jul-1999
C:Accession: D30502
R:Ellet, D.; Webster, D.M.; Rees, A.R.
J. Immunol. 141, 1745-1753, 1988
A:Title: V region sequences of anti-DNA and anti-RNA autoantibodies from NZB/NZW F-1 M.D.
A:Reference number: A30502; MUID:88315787
A:Accession: D30502
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-77 <EIL>
A:Cross-references: GB:M21908; NID:9197073; PIDN:AAA8908.1; PID:9197074
A>Note: the authors translated the codon CAG for residue 48 as Pro
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 32; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 40 FTLTSS 46

RESULT 4

S70443
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 07-May-1999
C:Accession: S70443
R:Cuisinier, A.M.; Fumoux, F.; Fougereau, M.; Tonnelie, C.
Mol. Immunol. 29, 1363-1373, 1992
A:Title: Igm kappa/lambda EBY human B cell clone: an early step of differentiation of B
A:Reference number: S70442; MUID:93024508
A:Accession: S70443
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-77 <CUU>
C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 100.0%; Score 32; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 40 FTLTSS 46

RESULT 5

S34102
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
C:Accession: S34102
R:Magner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34102
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-78 <MAG>
A:Cross-references: EMBL:X67186
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 32; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 62 FTLTSS 68

RESULT 6

S24215
Ig kappa chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24215
R:Shimizu, T.; Iwasato, T.; Yamagishi, H.
J. Exp. Med. 173, 1065-1072, 1991
A:Title: Deletions of immunoglobulin C(kappa) region characterized by the circular ex
A:Reference number: S24214; MUID:91217618
A:Accession: S24215
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-79 <SHI>
A:Cross-references: EMBL:X58202; NID:953718; PIDN:CAA41178.1; PID:930195
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 32; DB 2; Length 79;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 44 FTLTSS 50

RESULT 7

PH1048
Ig light chain V region (clone 165.49) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Jun-1996
C:Accession: PH1048
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both Igm and Igg anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1048
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-81 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 100.0%; Score 32; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTSS 7
|||||||
Db 56 FTLTSS 62

RESULT 8

S78489

Ig kappa chain V region (patient 28) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 21-Jan-2000
C:Accession: S78489; S34100
R:Wagner, S.
submitted to the EMBL Data Library, July 1992
A:Reference number: S78488
A:Accession: S78489
A:Molecule type: DNA
A:Residues: 1-83 <NAG>
A:Cross-references: EMBL:X67184
A:Experimental source: patient 28
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34100
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-68, 'Q', '70-73, 'A', '75-83 <NAG>
A:Cross-references: EMBL:X67184
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:6-80/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTIS 7
|||||
DB 61 FTLTIS 67

RESULT 9
S34099
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34099
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34099
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <NAG>
A:Cross-references: EMBL:X67183
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:5-79/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTIS 7
|||||
DB 60 FTLTIS 66

RESULT 10
S16826
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16826; S34101
R:Blaison, G.; Kuntz, J.L.; Pasquali, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid fad

A:Reference number: S16823; MUID:91243737
A:Accession: S16826
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54824; NID:q33653; PTDN:CAA38593.1; PID:q33654
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distrib
A:Reference number: S34076; MUID:93170387
A:Accession: S34101
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <NAG>
A:Cross-references: EMBL:X67185
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTIS 7
|||||
DB 62 FTLTIS 68

RESULT 11
S16834
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16834
R:Blaison, G.; Kuntz, J.L.; Pasquali, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid
A:Reference number: S16823; MUID:91243737
A:Accession: S16834
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54832
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTIS 7
|||||
DB 62 FTLTIS 68

RESULT 12
S16836
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16836
R:Blaison, G.; Kuntz, J.L.; Pasquali, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid
A:Reference number: S16823; MUID:91243737
A:Accession: S16836
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54834

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTISS 7
|||||
Db 62 FTLTISS 68

RESULT 13
S34086
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34086
R:Wagner, S.D.; Lanzetta, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34086
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-86 <KAG>
A:Cross-references: EMBL:X67169
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:9-83/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTISS 7
|||||
Db 64 FTLTISS 70

RESULT 14
S16840
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16840
R:Blaisson, G.; Kuntz, J.L.; Pasquall, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid fac
A:Reference number: S16823; MUID:91243737
A:Accession: S16840
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54838
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTISS 7
|||||
Db 62 FTLTISS 68

RESULT 13

S16837
Ig kappa chain V region - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S16837
R:Blaisson, G.; Kuntz, J.L.; Pasquall, J.L.
Eur. J. Immunol. 21, 1221-1227, 1991
A:Title: Molecular analysis of V(kappa)III variable regions of polyclonal rheumatoid
A:Reference number: S16823; MUID:91243737
A:Accession: S16837
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-86 <BLA>
A:Cross-references: EMBL:X54835
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-81/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 32; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTISS 7
|||||
Db 62 FTLTISS 68

Search completed: July 15, 2002, 13:00:57
Job time: 463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:59:24 ; Search time 44.9 Seconds
(without alignments)
6.036 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32
Sequence: 1 FTLLTSS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	107	KVID_HUMAN	P01596 homo sapien
2	32	100.0	108	KVIF_HUMAN	P01598 homo sapien
3	32	100.0	108	KVIG_HUMAN	P01599 homo sapien
4	32	100.0	108	KVIL_HUMAN	P01600 homo sapien
5	32	100.0	108	KVLS_HUMAN	P01611 homo sapien
6	32	100.0	108	KV5Q_MOUSE	P01650 mus musculu
7	32	100.0	108	KV5R_MOUSE	P01651 mus musculu
8	32	100.0	108	KV5S_MOUSE	P01652 mus musculu
9	32	100.0	108	KV5T_MOUSE	P01653 mus musculu
10	32	100.0	109	KVIF_HUMAN	P01612 homo sapien
11	32	100.0	109	KV3F_HUMAN	P01624 homo sapien
12	32	100.0	114	KVIA_MOUSE	P01632 homo sapien
13	32	100.0	114	KV4A_HUMAN	P01625 homo sapien
14	32	100.0	115	KV3I_HUMAN	P04433 homo sapien
15	32	100.0	116	KV3J_HUMAN	P04434 homo sapien
16	32	100.0	117	KV1I_HUMAN	P01601 homo sapien
17	32	100.0	117	KV1J_HUMAN	P01602 homo sapien
18	32	100.0	121	KV4O_HUMAN	P06312 homo sapien
19	32	100.0	129	KV1W_HUMAN	P04431 homo sapien
20	32	100.0	129	KV1X_HUMAN	P04432 homo sapien
21	32	100.0	133	KV4B_HUMAN	P06313 homo sapien
22	32	100.0	134	KV4C_HUMAN	P06314 homo sapien
23	32	100.0	136	KV5B_MOUSE	P01634 mus musculu
24	30	93.8	354	OPST_ASTRA	P51474 astyanax fa
25	29	90.6	108	KV1L_HUMAN	P01604 homo sapien
26	29	90.6	108	KV1Q_HUMAN	P01609 homo sapien
27	29	90.6	108	KV1R_HUMAN	P01610 homo sapien
28	29	90.6	400	CN37_BOVIN	P06623 bos taurus
29	29	90.6	421	CN37_HUMAN	P09543 homo sapien
30	28	87.5	92	KV09_RABIT	P01650 oryctolagus
31	28	87.5	100	KV3C_HUMAN	P01631 homo sapien
32	28	87.5	104	KV17_RABIT	P01658 oryctolagus
33	28	87.5	107	KV04_RABIT	P01685 oryctolagus

34	28	87.5	108	KV05_RABIT	P01686 oryctolagus
35	28	87.5	108	KV06_RABIT	P01687 oryctolagus
36	28	87.5	108	KV07_RABIT	P01688 oryctolagus
37	28	87.5	108	KV08_RABIT	P01689 oryctolagus
38	28	87.5	108	KV1B_HUMAN	P01594 homo sapien
39	28	87.5	108	KV1E_HUMAN	P01597 homo sapien
40	28	87.5	108	KV1G_HUMAN	P01603 homo sapien
41	28	87.5	108	KV1K_HUMAN	P01605 homo sapien
42	28	87.5	108	KV1M_HUMAN	P01606 homo sapien
43	28	87.5	108	KV1P_HUMAN	P01608 homo sapien
44	28	87.5	108	KV1Y_HUMAN	P00362 homo sapien
45	28	87.5	108	KV3A_HUMAN	P01619 homo sapien

ALIGNMENTS

```

RESULT 1
ID KVID_HUMAN          STANDARD:      PRT;      107 AA.
AC P01596;
DT 21-JUL-1986 (Rel. 01, Created)
DE 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region CAR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=75075135; PubMed=4216454;
RA Milstein C.P.; Deverson E.V.;
RT "Primary structure of kappa light chain from a human myeloma
RT protein."
RL Eur. J. Biochem. 49:377-391(1974).
RC -I- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
DR PIR: A01864; KIHUW.
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Glycoprotein.
FT CARBOHYD 28
FT NON_TER 107
SQ SEQUENCE      107 AA; 11703 MW; E1BF0DF9844C3346 CRC64;

Query Match          100.0%; Score 32; DB 1; Length 107;
Best local similarity 100.0%; Pred. No. 0.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLLTSS 7
Db 71 FTLLTSS 77

RESULT 2
ID KVIF_HUMAN          STANDARD:      PRT;      108 AA.
AC P01598;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region EU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE.
RX MEDLINE-71064023; PubMed-5489770;
RA Gottlieb P.D., Cunningham B.A., Rutishauser U., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. VI. Amino
RL acid sequence of the light chain."
RU Biochemistry 9:3155-3161(1970).
RN (2)
RP DISULFIDE BOND.
RX MEDLINE-71064027; PubMed-4923144;
RA Gali W.E., Edelman G.M.;
RT "The covalent structure of a human gamma G-immunoglobulin. X.
RL Intrachain disulfide bonds."
RU Biochemistry 9:3188-3196(1970).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC PIR: A01866; KIHUHU.
DR HSSP: P01607; IREI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON_TER 108
SQ SEQUENCE 108 AA; 11788 MW; 9CD294F2FD88823 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
DB 71 FTLTSS 77

RESULT 3
KV1G_HUMAN STANDARD; PRT; 108 AA.
AC P01599;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE-75059122; PubMed-4215718;
RA Laure C.J., Watanabe S., Hilschmann N.;
RT "The primary structure of a monoclonal Igm-immunoglobulin
RL (macroglobulin gal.), I. The amino acid sequence of the L-chain of
RT kappa-type, subgroup I."
RU Hoppe-Seyler's Z. Physiol. Chem. 354:1503-1504(1973).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
DR PIR: A01867; KIHUGL.
DR HSSP: P01607; IREI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.

```

```

FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON_TER 108
SQ SEQUENCE 108 AA; 11814 MW; C1AD3CB0F600F73 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
DB 71 FTLTSS 77

RESULT 4
KV1H_HUMAN STANDARD; PRT; 108 AA.
AC P01600;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Hau.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE-71032830; PubMed-4097974;
RA Watanabe S., Hilschmann N.;
RT "The primary structure of a monoclonal kappa-type immunoglobulin L-
RL chain of subgroup I (Bence-Jones protein Hau); subdivision within
RT subgroups."
RU Hoppe-Seyler's Z. Physiol. Chem. 351:1291-1295(1970).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR: A01868; KIHUHU.
DR HSSP: P80362; IWTU.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT DOMAIN 35 49
FT DOMAIN 50 56
FT DOMAIN 57 88
FT DOMAIN 89 97
FT DOMAIN 98 107
FT DISULFID 23 88
FT NON_TER 108
SQ SEQUENCE 108 AA; 11671 MW; 08D3A61608D0618 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
DB 71 FTLTSS 77

RESULT 5
KV1S_HUMAN

```

ID	KV15_HUMAN	STANDARD;	PRT;	108 AA.
AC	P01611;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig kappa chain V-I region Wes.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=81092279; PubMed=6778806;			
RA	Kratzlin H., Yang C., Krusche J.U., Hilschmann N.;			
RT	"Preparative separation of the tryptic hydrolysate of a protein by			
RT	high pressure liquid chromatography. The primary structure of a			
RT	monoclonal L-chain of k-type, subgroup I (Bence-Jones Protein			
RT	Wes).";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 361:1591-1598(1980).			
CC	-1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.			
CC	-1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.			
CC	PIR: A01877; KIHUMS.			
DR	HSSP; P80362; IWTL.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF00047; 1g_1.			
DR	SMART; SM00406; IgV_1.			
KW	Immunoglobulin V region; Bence-Jones protein.			
FT	DOMAIN	1	23	FRAMEWORK-1.
FT	DOMAIN	24	34	COMPLEMENTARITY-DETERMINING-1.
FT	DOMAIN	35	49	FRAMEWORK-2.
FT	DOMAIN	50	56	COMPLEMENTARITY-DETERMINING-2.
FT	DOMAIN	57	88	FRAMEWORK-3.
FT	DOMAIN	89	97	COMPLEMENTARITY-DETERMINING-3.
FT	DOMAIN	98	107	FRAMEWORK-4.
FT	DISULFD	23	88	BY SIMILARITY.
FT	NON_TER	108	108	
SQ	SEQUENCE	108 AA;	11608 MW;	782B14A649A60E45 CRC64;
Query Match				
Best Local Similarity		100.0%;	Score 32;	DB 1;
Matches		7;	Conservative	0;
			Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1 FTLTSS 7			
DB	71 FTLTSS 77			
RESULT 6				
KV5Q_MOUSE				
ID	KV5Q_MOUSE	STANDARD;	PRT;	108 AA.
AC	P01650;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig kappa chain V-V region UPC 61.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE.			
RX	MEDLINE=79195288; PubMed=109517;			
RA	Vrana M., Rudikoff S., Potter M.;			
RT	"The structural basis of a happen-inhitable kappa-chain idioctype.";			
RT	J. Immunol. 122:1905-1910(1979).			
CC	-1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT			
CC	BIND BEA(2-1)-FRUCTOFURANOSYL MOETIES (INULIN).			
CC	PIR: A01929; KYMS61.			
DR	HSSP; P80362; IWTL.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003596; Ig_V.			

```

DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11809 MW; 35C116BD60F79310 CRC64;

OY 1 FTLTSS 7 100.0%; Score 32; DB 1; Length 108;
    ||||| Best Local Similarity 100.0%; Pred. No. 0.81;
    71 FTLTSS 77 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 FTLTSS 77

RESULT 7
KVSR_MOUSE STANDARD; PRT; 108 AA.
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudnikoff S., Potter M.;
RT J. Immunol. 122:1905-1910(1979).
RL *- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA2-1-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B97808; KVM509.
DR HSSP: P01607; 1REX.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7 100.0%; Score 32; DB 1; Length 108;
    ||||| Best Local Similarity 100.0%; Pred. No. 0.81;
    71 FTLTSS 77 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 FTLTSS 77

RESULT 7
KVSR_MOUSE STANDARD; PRT; 108 AA.
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudnikoff S., Potter M.;
RT J. Immunol. 122:1905-1910(1979).
RL *- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA2-1-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B97808; KVM509.
DR HSSP: P01607; 1REX.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7 100.0%; Score 32; DB 1; Length 108;
    ||||| Best Local Similarity 100.0%; Pred. No. 0.81;
    71 FTLTSS 77 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 FTLTSS 77

RESULT 7
KVSR_MOUSE STANDARD; PRT; 108 AA.
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudnikoff S., Potter M.;
RT J. Immunol. 122:1905-1910(1979).
RL *- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA2-1-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B97808; KVM509.
DR HSSP: P01607; 1REX.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7 100.0%; Score 32; DB 1; Length 108;
    ||||| Best Local Similarity 100.0%; Pred. No. 0.81;
    71 FTLTSS 77 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 FTLTSS 77

RESULT 7
KVSR_MOUSE STANDARD; PRT; 108 AA.
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudnikoff S., Potter M.;
RT J. Immunol. 122:1905-1910(1979).
RL *- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA2-1-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B97808; KVM509.
DR HSSP: P01607; 1REX.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 108;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7 100.0%; Score 32; DB 1; Length 108;
    ||||| Best Local Similarity 100.0%; Pred. No. 0.81;
    71 FTLTSS 77 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 FTLTSS 77

RESULT 7
KVSR_MOUSE STANDARD; PRT; 108 AA.
AC P01651;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudnikoff S., Potter M.;
RT J. Immunol. 122:1905-1910(1979).
RL *- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA2-1-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B97808; KVM509.
DR HSSP: P01607; 1REX.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT 24 34 COMPLEMENTARITY-DETERMINING-1.
FT 35 49 FRAMEWORK-2.
FT 50 56 COMPLEMENTARITY-DETERMINING-2.
FT 57 88 FRAMEWORK-3.
FT 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 108;
Best Local Similarity 100.0
```

```

KV55_MOUSE
ID KV55_MOUSE STANDARD: PRT; 108 AA.
AC P01653;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Ig kappa chain V-V region J606.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma Inulin-
binding proteins.";
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
-I- BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: A92811; KVM506.
DR HSSP: P01607; 1REI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; IgV_1.
DR SMART: SM00406; IgV_1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108
SQ SEQUENCE 108 AA; 11810 MW; 8DE4DD31076F2A9B CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7
DB 71 FTLTSS 77

RESULT 9
KV5T_MOUSE
ID KV5T_MOUSE STANDARD: PRT; 108 AA.
AC P01653;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Ig kappa chain V-V region W3082.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma Inulin-
binding proteins.";
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
-I- BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B92811; KVM582.
DR HSSP: P80362; 1MTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.

```

```

DR Pfam: PF00047; IgV_1.
DR SMART: SM00406; IgV_1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 FRAMEWORK-4.
FT DOMAIN 98 108 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108
SQ SEQUENCE 108 AA; 11850 MW; C5C145DC376F30CD CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7
DB 71 FTLTSS 77

RESULT 10
KV1T_HUMAN
ID KV1T_HUMAN STANDARD: PRT; 109 AA.
AC P01612;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
OS Ig kappa chain V-I region Mev.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83081018; PubMed=6816713;
RA Eulitz M., Linke R.P.;
RT "Primary structure of the variable part of an amyloidogenic
Bence-Jones Protein (Mev). An unusual insertion in the third
hypervariable region of a human kappa-immunoglobulin light chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1347-1358(1982).
CC -I- MISCELLANEOUS: ANOTHER FORM THAT LACKED RESIDUES 1-3 WAS ALSO
FOUND.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.
DR PIR: A01879; KIH0WV.
DR HSSP: P01730; 1WIO.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; IgV_1.
DR SMART: SM00406; IgV_1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 98 FRAMEWORK-4.
FT DOMAIN 99 108 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 109
SQ SEQUENCE 109 AA; 11870 MW; B6ABF4515D5F5A0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 109;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7
DB 71 FTLTSS 77

```



```
RESULT 11
KV3F_HUMAN STANDARD: PRT; 109 AA.
AC P01624;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig kappa chain V-II region POM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76276460; PubMed=60899;
RA Klapper D.G., Capra J.D.;
RT "The amino acid sequence of the variable regions of the light chains
from two idiotypically cross reactive Igm anti-gamma globulins.";
RL Ann. Immunol. (Paris) 127C:261-271(1976).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01897; K3HUPM.
DR HSSP; P01789; 2MCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; Igv_1.
KW Immunoglobulin V region.
FT DISULFID 23
FT NON_TER 109
SQ SEQUENCE 109 AA; 11922 MW; 62821DDC6A8ABA86 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 109;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
Db 72 FTLTSS 78

RESULT 12
KV1A_MOUSE STANDARD: PRT; 114 AA.
AC P01632;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region S107A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81241357; PubMed=6788890;
RA Kwan S.-P., Rudikoff S., Seidman J.G., Leder P., Scharff M.D.;
RT "Nucleic acid and protein sequences of phosphocholine-binding light
chains.";
RL J. Exp. Med. 153:1366-1370(1981).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 029423; AAC00033.1; -.
```

```
DR PIR; A01915; KVNS7A.
DR HSSP; P01789; 1MCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region.
FT DISULFID 23
FT NON_TER 114
SQ SEQUENCE 114 AA; 12717 MW; 32008EC8E9DBE67B CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 114;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
Db 77 FTLTSS 83

RESULT 13
KV4A_HUMAN STANDARD: PRT; 114 AA.
AC P01625;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1986 (Rel. 34, Last sequence update)
DT 13-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region Ien.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76004342; PubMed=50995;
RA Schneider M., Hilschmann N.;
RT "The primary structure of a monoclonic immunoglobulin-L-chain of
subgroup IV of the kappa type (Bence-Jones protein Ien).";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN [2]
RP REVISION TO 9.
RA Salomon A.;
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR; A01903; KAHULN.
DR HSSP; P01789; 1MCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DISULFID 23
FT NON_TER 114
SQ SEQUENCE 114 AA; 12640 MW; 0647F1D17F236485 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 114;
```

Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FTLTSS 7
Db 77 FTLTSS 83

RESULT 14
KV3J_HUMAN STANDARD; PRT; 115 AA.
AC P04433;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region VG precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087932; PubMed=6440122;
RA Pech M., Zachau H.G.;
RT Immunoglobulin genes of different subgroups are interdigitated
RT within the VK locus*;
RL Nucleic Acids Res. 12:9229-9236(1984).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL; X01668; -; NOT_ANNOTATED_CDS.
DR PIR: A01900; K3HUVG.
DR HSSP: P01789; 1MCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
KW SIGNAL 1 20
FT CHAIN 21 >115 IG KAPPA CHAIN V-II REGION VG.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 69 FRAMEWORK-2.
FT DOMAIN 70 76 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 77 108 FRAMEWORK-3.
FT DOMAIN 109 115 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 108 BY SIMILARITY.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12575 MW; 2DE47CDA3A17D555 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7
Db 91 FTLTSS 97

RESULT 15
KV3J_HUMAN STANDARD; PRT; 116 AA.
AC P04434;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig kappa chain V-II region VH precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85087932; PubMed=6440122;
RA Pech M., Zachau H.G.;
RT Immunoglobulin genes of different subgroups are interdigitated
RT within the VK locus*;
RL Nucleic Acids Res. 12:9229-9236(1984).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
CC EMBL; X02725; -; NOT_ANNOTATED_CDS.
DR PIR: A01901; K3HUVH.
DR HSSP: P01789; 2MCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
KW SIGNAL 1 20
FT CHAIN 21 >116 IG KAPPA CHAIN V-II REGION VH.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 55 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 56 70 FRAMEWORK-2.
FT DOMAIN 71 77 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 78 109 FRAMEWORK-3.
FT DOMAIN 110 116 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 109 BY SIMILARITY.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12757 MW; 51CD55BA53B21929 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.88;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLTSS 7
Db 92 FTLTSS 98

Search completed: July 15, 2002, 13:23:30
Job time: 1446 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:58:04 ; Search time 172.49 Seconds

(without alignments)
7.020 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32

Sequence: 1 FTLTISS 7

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP rvirus:*
16: SP bacteriophage:*
17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	107	4	Q96SA9
2	32	100.0	107	11	Q9BRZ9
3	32	100.0	108	4	Q9UL83
4	32	100.0	108	4	Q9UL77
5	32	100.0	108	4	Q9UL70
6	32	100.0	109	4	Q9UL85
7	30	93.8	1004	17	Q28412
8	29	90.6	116	4	Q96PF6
9	28	90.6	171	12	Q6S504
10	29	90.6	195	12	Q9J0T0
11	29	90.6	195	12	Q9J0S8
12	29	90.6	196	12	Q9J0T8
13	29	90.6	196	12	Q9J0T7
14	29	90.6	196	12	Q9J0T2
15	29	90.6	203	12	Q9J0T6
16	29	90.6	203	12	Q9J0T4

17	29	90.6	204	12	Q9J0T5
18	29	90.6	205	12	Q9J0S9
19	29	90.6	206	12	Q9J0S7
20	29	90.6	206	12	Q9J0S6
21	29	90.6	206	12	Q9J0S5
22	29	90.6	206	12	Q9J0S4
23	29	90.6	206	12	Q9J0S3
24	29	90.6	206	12	Q9J0S2
25	29	90.6	206	12	Q9J0S0
26	29	90.6	206	12	Q9J0S9
27	29	90.6	206	12	Q9J0S8
28	29	90.6	206	12	Q9J0S7
29	29	90.6	206	12	Q9J0S6
30	29	90.6	206	12	Q9J0S5
31	29	90.6	206	12	Q9J0S4
32	29	90.6	206	12	Q9J0S3
33	29	90.6	206	12	Q9J0S2
34	29	90.6	206	12	Q9J0S1
35	29	90.6	206	12	Q9J0S0
36	29	90.6	206	12	Q9J0S9
37	29	90.6	206	12	Q9J0S8
38	29	90.6	206	12	Q9J0S7
39	29	90.6	206	12	Q9J0S6
40	29	90.6	206	12	Q9J0S5
41	29	90.6	206	12	Q9J0S4
42	29	90.6	206	12	Q9J0S3
43	29	90.6	206	12	Q9J0S2
44	29	90.6	206	12	Q9J0S1
45	29	90.6	206	12	Q9J0S0

ALIGNMENTS

RESULT 1					
Q96SA9	PRELIMINARY;	PRT;	107 AA.		
AC	Q96SA9;				
DT	01-DEC-2001 (TREMblrel. 19, Created)				
DT	01-DEC-2001 (TREMblrel. 19, last sequence update)				
DE	ANTI-STREPTOCOCCAL/ANTI-MYOSIN IMMUNOGLOBULIN KAPPA LIGHT CHAIN				
DE	VARIABLE REGION (FRAGMENT).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid-9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-98375893; PubMed-9712075;				
RA	Adderson E.E., Shikhan A.R., Ward K.E., Cunningham M.W.;				
RT	*Molecular analysis of polyclonal antibodies from				
RT	rheumatic carditis: human anti-N-acetylglucosamine/anti-myosin				
RT	antibody V region genes.;				
RL	J. Immunol. 161:2020-2031(1998).				
DR	EMBL; U96396; AAB6785.1; -.				
FT	NON_TER	1			
FT	NON_TER	107			
SO	SEQUENCE	107 AA.	11520 MW;	4BB43E9C5B577F16	CRC64;

Query Match 100.0%; Score 32; DB 4; Length 107;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 FTLTISS 7				
DB	71 FTLTISS 77				
RESULT 2					
Q9BRZ9	PRELIMINARY;	PRT;	107 AA.		
ID	Q9BRZ9				

AC 09ER29;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ANTI HUMAN TNF-ALPHA LIGHT CHAIN VARIABLE REGION (FRAGMENT).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
 RT Cloning and sequencing of the light chain fragment of variable region
 RT genes of an anti-TNF- α monoclonal antibody.";
 RL J. Cell. Mol. Immunol. 12:21-26(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Su C.Z.;
 RT Construction and sequencing of the single-chain antibody gene of a
 RT human TNF- α specific monoclonal antibody.";
 RL 114 Chun 1 Ta Hsueh Hsueh Pao 19:373-376(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen P., Deng J.B., Wang Z.L., Han H., Yao L.B., Su C.Z.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF262753; AAG23804.1; -.
 DR HSSP; P80362; 1WTL.
 DR InterPro: IPR003599; I9.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003596; I9_V.
 DR Pfam; PF00047; I9; 1.
 DR SMART; SM00409; I9; 1.
 DR SMART; SM00406; IGV; 1.
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

Query Match 100.0%; Score 32; DB 11; Length 107;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
 |||||
 DB 74 FTLTSS 80

RESULT 3
 09UL83 PRELIMINARY; PRT; 108 AA.
 AC 09UL83;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
 DE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 OC Clin. Immunol. Immunopathol. 87:184-192(1999).
 DR EMBL; AF035031; AAD56267.1; -.
 DR HSSP; P80362; 1WTL.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003596; I9_V.
 DR Pfam; PF00047; I9; 1.
 DR SMART; SM00406; IGV; 1.
 FT NON_TER 107 107
 SQ SEQUENCE 107 AA; 11784 MW; 2B15EEA6604A26C3 CRC64;

FT NON_TER 1 1
 FT NON_TER 108 108
 SQ SEQUENCE 108 AA; 11834 MW; 9F9C5A92BBA96EEA CRC64;

Query Match 100.0%; Score 32; DB 4; Length 108;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
 |||||
 DB 71 FTLTSS 77

RESULT 4
 09UL77 PRELIMINARY; PRT; 108 AA.
 AC 09UL77;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
 DE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 RT fetus".
 OC Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL; AF035031; AAD56273.1; -.
 DR HSSP; P01607; 1REI.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR003596; I9_V.
 DR Pfam; PF00047; I9; 1.
 DR SMART; SM00406; IGV; 1.
 FT NON_TER 1 1
 FT NON_TER 108 108
 SQ SEQUENCE 108 AA; 11738 MW; C06681716C4D16F3 CRC64;

Query Match 100.0%; Score 32; DB 4; Length 108;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
 |||||
 DB 71 FTLTSS 77

RESULT 5
 09UL70 PRELIMINARY; PRT; 108 AA.
 AC 09UL70;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
 DE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98277139; PubMed-9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
 RA Young D.C.;

```

RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035044; AAD56280.1; -.
DR HSSP; P01607; IREI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT SEQUENCE 108 AA; 11653 MW; B7BEDC3E41FCCA37 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 108;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTUTISS 7
DB 71 FTUTISS 77

RESULT 6
Q9UL85 PRELIMINARY; PRT; 109 AA.
AC Q9UL85;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN KAPPA CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035029; AAD56265.1; -.
DR HSSP; P80362; IPTL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT SEQUENCE 109 AA; 11761 MW; FB1E43E7C7AFACCC CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 109;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTUTISS 7
DB 71 FTUTISS 77

RESULT 7
Q28412 PRELIMINARY; PRT; 1004 AA.
AC Q28412;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)
DE HYPOTHETICAL 115.0 KDA PROTEIN.
GN AF1867.

```

```

OS Archaeoglobus fulgidus.
OC Archaeae; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleuk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwyn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RA "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE000974; AAB89390.1; -.
DR TIGR; AF1867; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1004 AA; 115036 MW; 10DB5A7A3E900F6B CRC64;

Query Match
Best Local Similarity 93.8%; Score 30; DB 17; Length 1004;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTUTISS 7
DB 402 FTUTISS 408

RESULT 8
Q96PF6 PRELIMINARY; PRT; 116 AA.
AC Q96PF6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KAPPA 1 LIGHT CHAIN VARIABLE REGION (FRAGMENT).
GN SDMK1
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21361171; PubMed=11468171;
RA Comenzo R.L., Zhang Y., Martinez C., Osman K., Herrera G.A.;
RT "The tropism of organ involvement in primary systemic amyloidosis:
RT contributions of Ig V(L) germ line gene use and clonal plasma cell
RT burden.";
RL Blood 98:714-720(2001).
DR EMBL; AF361758; AAK51465.1; -.
FT NON_TER 1
FT SEQUENCE 116 AA; 12735 MW; E796FC2217BFCF57 CRC64;

Query Match
Best Local Similarity 90.6%; Score 29; DB 4; Length 116;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTUTISS 7
DB 71 FTUTISS 77

```

```
RESULT 9
ID 069504 PRELIMINARY; PRT; 171 AA.
AC 069504;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-1998 (TREMBLrel. 06, Last annotation update)
DE U23 PROTEIN.
GN U23.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=J1;
RA Nichols J.;
RL Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K;
RA Megaw A.G., Rapaport D., Avidor B., Frenkel N., Davison A.J.;
RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K;
RA Megaw A.G., Frenkel N., Davison A.J.;
RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF034218; AAC40736.1; -.
DR EMBL; AF034218; AAC40736.1; -.
SQ SEQUENCE 171 AA; 18663 MW; ADE85FA5909D8C20 CRC64;
```

```
Query Match
Best Local Similarity 90.6%; Score 29; DB 12; Length 171;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 FTLTITS 7
Db 98 FTLTIST 104
```

```
RESULT 10
ID 09J070 PRELIMINARY; PRT; 195 AA.
AC 09J070;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN 1 (FRAGMENT).
GN CPl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Seguiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PC4111;
RA Azzam O.I., Yambo M.L.M., Muhsin M., McNally K.L., Umahay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RT provinces of the Philippines and Indonesia.";
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223078; AAF65289.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 195 AA; 20789 MW; 0EA7DD0705357C7 CRC64;
```

```
Query Match
Best Local Similarity 90.6%; Score 29; DB 12; Length 195;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 FTLTITS 7
```

```
Db 74 FTLTITS 80

RESULT 11
ID 09J058 PRELIMINARY; PRT; 195 AA.
AC 09J058;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN 1 (FRAGMENT).
GN CPl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Seguiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PC8811;
RA Azzam O.I., Yambo M.L.M., Muhsin M., McNally K.L., Umahay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RT provinces of the Philippines and Indonesia.";
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223080; AAF65291.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 195 AA; 20865 MW; 11B3D6D0704450D7 CRC64;
```

```
Query Match
Best Local Similarity 90.6%; Score 29; DB 12; Length 195;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 FTLTITS 7
Db 74 FTLTITS 80
```

```
RESULT 12
ID 09J078 PRELIMINARY; PRT; 196 AA.
AC 09J078;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE COAT PROTEIN 1 (FRAGMENT).
GN CPl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Seguiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TB4111;
RA Azzam O.I., Yambo M.L.M., Muhsin M., McNally K.L., Umahay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RT provinces of the Philippines and Indonesia.";
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223070; AAF65281.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 196 AA; 20917 MW; CF8299B7E3D1F6A0 CRC64;
```

```
Query Match
Best Local Similarity 90.6%; Score 29; DB 12; Length 196;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 FTLTITS 7
Db 75 FTLTITS 81
```

```

RESULT 13
Q9J0T7 PRELIMINARY; PRT; 196 AA.
ID 09J0T7
AC 09J0T7;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE COAT PROTEIN 1 (FRAGMENT).
GN Cpl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sequiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IB4211;
RA Azam O.I., Yambao M.L.M., Muhsin M., McNally K.L., Umadhay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223071; AAF65282.1; -.
FT NON_TER 1 1
SQ SEQUENCE 196 AA; 20917 MW; CF8299B7E3D1F6A0 CRC64;

```

```

Query Match          90.6%; Score 29; DB 12; Length 196;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FTLTIS 7
DB 75 FTLTIS 81

```

```

RESULT 14
Q9J0T2 PRELIMINARY; PRT; 196 AA.
ID 09J0T2;
AC 09J0T2;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE COAT PROTEIN 1 (FRAGMENT).
GN Cpl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sequiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PC201;
RA Azam O.I., Yambao M.L.M., Muhsin M., McNally K.L., Umadhay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223076; AAF65287.1; -.
FT NON_TER 1 1
SQ SEQUENCE 196 AA; 20961 MW; 1E128B361CD743D CRC64;

```

```

Query Match          90.6%; Score 29; DB 12; Length 196;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FTLTIS 7
DB 75 FTLTIS 81

```

```

ID 09J0T6 PRELIMINARY; PRT; 203 AA.
AC 09J0T6;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE COAT PROTEIN 1 (FRAGMENT).
GN Cpl.
OS Rice tungro spherical virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Sequiviridae;
OC Waikavirus.
OX NCBI_TaxID=35287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PC03VI;
RA Azam O.I., Yambao M.L.M., Muhsin M., McNally K.L., Umadhay K.M.L.;
RT "Genetic diversity of rice tungro spherical virus in tungro-endemic
RL Arch. Virol. 145:0-0(2000).
DR EMBL; AF223072; AAF65283.1; -.
FT NON_TER 1 1
SQ SEQUENCE 203 AA; 21647 MW; E8CC8A6A3A1E1722 CRC64;

```

```

Query Match          90.6%; Score 29; DB 12; Length 203;
Best Local Similarity 85.7%; Pred. No. 48;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 FTLTIS 7
DB 82 FTLTIS 88

```

Search completed: July 15, 2002, 13:22:39
Job time: 1475 sec

RESULT 15
Q9J0T6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:51:04 : Search time 228.39 Seconds
(without alignments)
3.404 Million cell updates/sec

Title: US-09-712-819A-1

Perfect score: 32
Sequence: 1 FTUITS 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A: Geneseq 032802: *
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	32	100.0	32	17	AA887043	Human group I 11gh
2	32	100.0	32	20	AA552745	Humanised Atr-5 L
3	32	100.0	32	22	AA898286	Anti-A33 antigen i
4	32	100.0	32	22	AA898288	Anti-A33 antigen i
5	32	100.0	32	22	AA897666	A33 antigen bindin
6	32	100.0	32	22	AA897668	A33 antigen bindin
7	32	100.0	74	15	AA862921	Human cytomagalovi
8	32	100.0	74	19	AA862805	Human cytomagalovi
9	32	100.0	75	15	AA862923	Human cytomagalovi
10	32	100.0	76	20	AA80981	Variable kappa 1lg
11	32	100.0	82	19	AA862807	Amino acid sequenc

12	32	100.0	84	17	AAW14491	Monoclonal antibod
13	32	100.0	84	17	AA89878	Monoclonal antibod
14	32	100.0	86	19	AA862806	Amino acid sequenc
15	32	100.0	88	19	AAW5619	Anti-RSV F protein
16	32	100.0	88	21	AA156654	Partial peptide fr
17	32	100.0	88	21	AA156655	Partial peptide fr
18	32	100.0	88	21	AA156656	Partial peptide fr
19	32	100.0	88	21	AA156657	Partial peptide fr
20	32	100.0	88	21	AA156658	Partial peptide fr
21	32	100.0	88	21	AA156659	Partial peptide fr
22	32	100.0	88	21	AA156672	Partial peptide fr
23	32	100.0	94	19	AAW62808	Amino acid sequenc
24	32	100.0	94	22	AA867510	Light chain variab
25	32	100.0	95	13	AA828741	Human K102 VL. Ho
26	32	100.0	95	15	AA857478	OF7K.11 VK-1 L cha
27	32	100.0	95	16	AA872061	OF7K.9 VK-1 L cha
28	32	100.0	95	16	AA872062	OF7K.19 VK-1 L cha
29	32	100.0	95	16	AA872063	OF7K.17 VK-1 L cha
30	32	100.0	95	16	AA872064	S43434 VK-1 reglon
31	32	100.0	95	16	AA872065	K1012 VK reglon.
32	32	100.0	95	16	AA872058	OF7K.3 VK-1 L cha
33	32	100.0	95	16	AA872059	OF7K.16 VK-1 L cha
34	32	100.0	95	16	AA872060	Human K102 VL regl
35	32	100.0	95	17	AA892081	Human autoantibody
36	32	100.0	98	22	AA880218	Igg antibody 2.1.1
37	32	100.0	101	20	AA134316	Humanized 399 Ligh
38	32	100.0	102	21	AA156683	Anti-HIV gp120 imm
39	32	100.0	103	22	AA893599	Anti-HIV gp120 imm
40	32	100.0	104	15	AA854318	VL region of HIV n
41	32	100.0	104	17	AAW01285	Amino acid sequenc
42	32	100.0	104	19	AAW62803	Anti-gp120 antibod
43	32	100.0	104	21	AA15137	Anti-gp120 antibod
44	32	100.0	104	21	AA158246	Anti-HIV gp120 imm
45	32	100.0	105	15	AA854310	

ALIGNMENTS

RESULT 1
AA87043 standard; Peptide; 32 AA.
XX
AC AA87043;
XX
DT 25-JUN-1996 (first entry)
XX
DE Human group I light chain framework 3.
XX
KW Humanised antibody; interleukin-5; IL-5; recombinant antibody;
KW antibody engineering; monoclonal antibody; MAAb; 39D10; CDR;
KW complementarity determining region; light chain; framework;
KW eosinophilia; allergy; asthma.
XX
OS Homo sapiens.
XX
PN WO9535375-A1.
XX
PD 28-DEC-1995.
XX
PF 16-JUN-1995; 95WO-GB01411.
XX
PR 17-JUN-1994; 94GB-0012230.
XX
PA (CLUT) CELLTECH THERAPEUTICS LTD.
XX
PI Athwal DS, Bodmer MW, Emlage JS;
XX
DR WPI; 1996-058412/06.
XX
PT Anti-human IL-5 recombinant antibody - useful for preventing or
PT reducing eosinophilia and for treating certain allergic diseases,
PT esp. asthma

XX Example 3; Fig 3; 69pp; English.
PS Framework regions (AAR87041-44) of human group I (gpl) germ line
CC antibody light chain showed homology to corresponding regions
CC (AAR87045-48, respectively) of the rat anti-human interleukin-5
CC monoclonal antibody 39D10 light chain (see AAR87040). This homology
CC was utilised in the prodn. of a humanised 39D10 VL (AAR87057) in
CC which rat 39D10 VL complementarily determining regions were grafted
CC into the human gpl framework.
XX
SQ Sequence 32 AA:

Query Match 100.0%; Score 32; DB 17; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
 |||||
DB 15 flltiss 21

RESULT 2
AAV52745
ID AAV52745 standard; Peptide; 32 AA.
XX
XX AAV52745;
AC
XX
XX 26-JAN-2000 (first entry)
DT
XX
XX Humanised ATR-5 L chain V region FR3 for "a".
DE
XX
XX Human tissue factor; TF; humanised; antibody; mouse monoclonal antibody;
KW ATR-2; ATR-3; ATR-4; ATR-5; ATR-7; ATR-8; thrombotic disease; DIC;
KW disseminated intravascular coagulation; immunogenicity; chimeric.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO9951743-A1.
PN
XX 14-OCT-1999.
PD
XX 02-APR-1999; 99WO-JP01768.
PP
XX 03-APR-1998; 98JP-0091850.
PR
XX (CHUS) CHUGAI SEIYAKU KK.
PA
XX Sato K, Adachi H, Yabuta N;
PI
XX WPI; 1999-620204/53.
DR
XX
XX Humanised antibody recognizing human tissue factor, used for treatment
PT of disseminated intravascular coagulation -
XX
XX Claim 17; Page 270; 291pp; Japanese.
PS
XX The present invention describes chimeric antibody (Ab) heavy (H) chains
CC containing the variable region of the H chain of a mouse monoclonal Ab
CC recognising human tissue factor (htf) and the constant region of the H
CC chain of a human Ab. The variable region is one of six specified
CC sequences (which are the H chain variable regions from mouse monoclonal
CC Ab's ATR-2,3,4,5,7 or 8). Also described are chimeric Ab light (L)
CC chains containing the variable region of the L chain of a mouse
CC monoclonal Ab recognising human tissue factor (htf) and the constant
CC region of the L chain of a human Ab, the variable region being one of six
CC specified sequences (which are the L chain variable regions from mouse
CC monoclonal Ab's ATR-2,3,4,5,7 or 8). The chimeric Ab's can be used for
CC the treatment and prevention of thrombotic disease, especially of
CC disseminated intravascular coagulation (DIC). The humanised antibody has
CC the high htf binding activity of the mouse monoclonal antibody but

CC greatly reduced immunogenicity. AA233001 to AA233091 and Y527007 to
CC AAV52767 represent sequences used in the exemplification of the present
CC invention.
XX
XX SQ Sequence 32 AA:

Query Match 100.0%; Score 32; DB 20; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
 |||||
DB 15 flltiss 21

RESULT 3
AAB98286
ID AAB98286 standard; Peptide; 32 AA.
XX
XX AAB98286;
AC
XX
XX 20-AUG-2001 (first entry)
DT
XX
XX Anti-A33 antigen immunoglobulin VL FR3 SEQ ID NO:92.
DE
XX
XX Human; rabbit; humanised; A33 antigen; anti-A33 antigen antibody;
KW immunoglobulin; complementarily determining region; CDR; cancer;
KW cytostatic; anticancer; colon cancer; stomach cancer.
XX
XX Homo sapiens.
OS
XX WO200130393-A2.
PN
XX 03-MAY-2001.
PD
XX 20-OCT-2000; 2000WO-US29289.
PP
XX 22-OCT-1999; 99US-0425638.
PR
XX 04-APR-2000; 2000US-0543004.
PR
XX (LUDWIG) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
PA (SCRI) SCRIPPS RES INST.
XX
XX Barbas CF, Rader C, Rittner G, Welt S, Old LJ;
PI
XX WPI; 2001-328613/34.
DR
XX
XX Treating cancers, particularly of stomach and colon, that express A33
PT antigen by administering conjugate of anticancer agent with specific
PT immunoglobulin product -
XX
XX Claim 16; Page 40; 85pp; English.
PS
XX The present invention describes a method for treating cancers that
CC express the A33 antigen. The method comprises administering an
CC anticancer agent (I) conjugated to an immunoglobulin product (II) that
CC binds specifically to A33 and contains one or more of 13 specified
CC complementarily determining regions (CDRs), given in AAB98262 to
CC AAB98274. (I) has cytostatic activity. The method can be used for
CC treating colon and stomach cancers. (II), or the nucleic acid encoding
CC it, can be used directly, in unconjugated form, for immunotherapy of
CC cancer, and, when labeled, for detection or diagnosis of diseases
CC associated with A33 expression. AA422218 to AA422254 and AAB98230 to
CC AAB98321 represent sequences used in the exemplification of the
CC present invention.
XX
XX SQ Sequence 32 AA;

Query Match 100.0%; Score 32; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLTSS 7
|||||||

DB 15 fcltss 21

RESULT 4

AAB98288

ID AAB98288 standard; Peptide; 32 AA.

AC AAB98288;

DT 20-AUG-2001 (first entry)

DE Anti-A33 antigen immunoglobulin VL FR3 SEQ ID NO:94.

KW Human; rabbit; humanised; A33 antigen; anti-A33 antigen antibody;

KW immunoglobulin; complementarity determining region; CDR; cancer;

KW cytostatic; anticancer; colon cancer; stomach cancer.

OS Homo sapiens.

PN WO200130393-A2.

PD 03-MAY-2001.

PF 20-OCT-2000; 2000WO-US29289.

PR 22-OCT-1999; 99US-0425638.

PR 04-APR-2000; 2000US-0543004.

PA (LUDWIG-) LUDWIG INST CANCER RES.

PA (SLOK) SLOAN KETTERING INST CANCER RES.

PA (SCRI) SCRIPPS RES INST.

PI Barbas CF, Rader C, Ritter G, Welt S, Old LJ;

PI MPI: 2001-328613/34.

PT Treating cancers, particularly of stomach and colon, that express A33

PT antigen by administering conjugate of anticancer agent with specific

PT immunoglobulin product -

PS Claim 16; Page 40; 85pp; English.

CC The present invention describes a method for treating cancers that

CC express the A33 antigen. The method comprises administering an

CC anticancer agent (I) conjugated to an immunoglobulin product (II) that

CC binds specifically to A33 and contains one or more of 13 specified

CC complementarity determining regions (CDRs), given in AAB98286 to

CC AAB8274. (I) has cytostatic activity. The method can be used for

CC treating colon and stomach cancers. (II), or the nucleic acid encoding

CC it, can be used directly, in unconjugated form, for immunotherapy of

CC cancer, and, when labeled, for detection or diagnosis of diseases

CC associated with A33 expression. AAB22218 to AAB22254 and AAB98230 to

CC AAB98321 represent sequences used in the exemplification of the

CC present invention.

SQ Sequence 32 AA;

Query Match 100.0%; Score 32; DB 22; Length 32;

Best Local Similarity 100.0%; Pred. No. 3.2; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0;

QY 1 FTLTSS 7
|||||||

DB 15 fcltss 21

RESULT 5
AAB97666

ID AAB97666 standard; Peptide; 32 AA.

AC AAB97666;

DT 08-AUG-2001 (first entry)

DE A33 antigen binding immunoglobulin product VLFR3 peptide SEQ ID NO:92.

KW Chimeric antibody; humanised; humanisation; mammalian antibody; antigen;

KW immunoreact; anti A33 antigen antibody; immunoglobulin.

OS Homo sapiens.

PN WO200131065-A1.

PD 03-MAY-2001.

PF 20-OCT-2000; 2000WO-US29026.

PR 22-OCT-1999; 99US-0425638.

PR 04-APR-2000; 2000US-0543004.

PA (SCRI) SCRIPPS RES INST.

PI Barbas CF, Rader C;

PI MPI: 2001-328657/34.

PT Preparing humanized rabbit antibodies that specifically immunoreact

PT with a particular antigen using display technology for expressing

PT libraries of antibody domains and fine tuning variable domain regions -

PS Example 9; Page 39; 62pp; English.

CC The present invention describes a method for preparing a humanised rabbit

CC antibody that specifically immunoreacts with a particular antigen. The

CC method comprises expressing a library of antibodies comprising one or

CC more complementarity determining region (CDR) from the variable domain

CC sequences that specifically immunoreact with the antigen grafted into

CC framework regions from humans, and selecting the antibodies that react

CC with the antigen. The method is useful for humanising non-human

CC mammalian antibodies, which can be used for the treatment of a variety

CC of diseases. The present sequence represents an A33 antigen binding

CC immunoglobulin product VLFR3 peptide which is given in an example from

CC the present invention.

SQ Sequence 32 AA;

Query Match 100.0%; Score 32; DB 22; Length 32;

Best Local Similarity 100.0%; Pred. No. 3.2; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0;

QY 1 FTLTSS 7
|||||||

DB 15 fcltss 21

RESULT 6
AAB97668

ID AAB97668 standard; Peptide; 32 AA.

AC AAB97668;

DT 08-AUG-2001 (first entry)

DE A33 antigen binding immunoglobulin product VLFR3 peptide SEQ ID NO:94.

KW Chimeric antibody; humanised; humanisation; mammalian antibody; antigen;

KW immunoreact; anti A33 antigen antibody; immunoglobulin.

OS Homo sapiens.

PN WO200131065-A1.
 XX
 PD 03-MAY-2001.
 XX
 PF 20-OCT-2000; 2000WO-US29026.
 XX
 PR 22-OCT-1999; 99US-0425638.
 PR 04-APR-2000; 2000US-0543604.
 XX
 PA (SCRI) SCRIPPS RES INST.
 PI Barbas CF, Rader C;
 XX
 DR WPI; 2001-328657/34.
 XX
 PT Preparing humanized rabbit antibodies that specifically immunoreact
 PT with a particular antigen using display technology for expressing
 PT libraries of antibody domains and fine tuning variable domain regions -
 XX
 PS Example 9; Page 39; 62pp; English.
 PS
 XX
 CC The present invention describes a method for preparing a humanised rabbit
 CC antibody that specifically immunoreacts with a particular antigen. The
 CC method comprises expressing a library of antibodies comprising one or
 CC more complementarity determining region (CDR) from the variable domain
 CC sequences that specifically immunoreact with the antigen grafted into
 CC framework regions from humans, and selecting the antibodies that react
 CC with the antigen. The method is useful for humanising non-human
 CC mammalian antibodies, which can be used for the treatment of a variety
 CC of diseases. The present sequence represents an A33 antigen binding
 CC immunoglobulin product VLFR3 peptide which is given in an example from
 CC the present invention.
 CC
 SQ Sequence 32 AA:
 XX

Query Match 100.0%; Score 32; DB 22; Length 32;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FTLTSS 7
 |||||
 Db 15 fcltiss 21

RESULT 7
 AAR62921
 ID AAR62921 standard; protein; 74 AA.
 XX
 AC AAR62921;
 XX
 DT 18-JUL-1995 (first entry)
 DT
 XX
 DE Human cytomegalovirus antibody light chain variable region.
 DE
 XX
 KM Human cytomegalovirus; antibody light chain variable region;
 KM HCMV; Immunocassay; Immunotherapy.
 KM
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO9425490-A.
 PN
 XX
 PD 10-NOV-1994.
 PD
 XX
 PF 29-APR-1994; 94WO-US04705.
 PF
 XX
 PR 30-APR-1993; 93US-0055985.
 PR
 XX
 PA (SCRI) SCRIPPS RES INST.
 PA
 XX
 PI Barbas C, Burton R, Burton DR, Williamson A, Burtonl R;
 PI
 XX
 DR WPI; 1994-358194/44.
 DR

XX
 XX Human monoclonal antibodies (Mabs) against human cytomegalovirus
 PT - also nucleic acids and cell lines producing the Mabs, useful in
 PT diagnosis and immunotherapy
 PT
 XX
 PS Claim 6; Page 150; 171pp; English.
 PS
 XX
 CC AAR62618-R62920 and AAR62921-R62923 are human cytomegalovirus (HCMV)
 CC antibody heavy chain and light chain variable regions respectively.
 CC A monoclonal antibody (Mab) containing a pair of these regions can be
 CC used to detect HCMV, and anti-human HCMV Abs in human patients via
 CC a competitive immunoassay. The Mabs may also be useful in
 CC immunotherapy.
 CC
 SQ Sequence 74 AA:
 XX

Query Match 100.0%; Score 32; DB 15; Length 74;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 FTLTSS 7
 |||||
 Db 34 fcltiss 40

RESULT 8
 AAM62805
 ID AAM62805 standard; Peptide; 74 AA.
 XX
 AC AAM62805;
 XX
 DT 23-SEP-1998 (first entry)
 DT
 XX
 DE Amino acid sequence of a human antibody fragment.
 DE
 XX
 KM Human; immunoglobulin; Ig; transgenic; non-human mammal;
 KM inactivated endogenous Ig locus; B-cell development;
 KM human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
 KM kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
 KM production; antibody.
 KM
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO9824893-A2.
 PN
 XX
 PD 11-JUN-1998.
 PD
 XX
 PF 03-DEC-1997; 97WO-US23091.
 PF
 XX
 PR 03-DEC-1996; 96US-0759620.
 PR
 XX
 PA (ABGE-) ABGENIX INC.
 PA
 XX
 PI Green L, Jakobovits A, Klapholz S, Kuchertlapatl R;
 PI Mendez M;
 PI
 XX
 DR WPI; 1998-333314/29.
 DR
 XX
 PT New transgenic non-human mammals - having an inactivated
 PT immunoglobulin locus and a near complete human immunoglobulin locus,
 PT used for production of human antibodies
 PT
 XX
 PS Disclosure; Page 77; 128pp; English.
 PS
 XX
 CC AAM62793-822 represent fragments of human antibodies produced by
 CC transgenic Xenomice, created using the method of the invention. The
 CC specific genetic Xenomice, created using the method of the invention. The
 CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
 CC locus, so that the mammal does not display normal B-cell development. The
 CC modified genome also has an inserted human heavy chain Ig locus in
 CC germ-line configuration, the human heavy chain Ig locus comprising a human
 CC micro constant region and regulatory and switch sequences, human J-H

CC genes, human D-H genes, and human V-H genes and an inserted human kappa
 CC light chain Ig locus in germline configuration, the human kappa light
 CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
 CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
 CC are selected to restore normal B-cell development in the mammal. The
 CC transgenic animals have a near complete human Ig locus, including both a
 CC human heavy chain locus and a human kappa light chain locus. They can
 CC be used for the production of human antibodies when exposed to
 CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
 CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
 CC respectively.

XX Sequence 74 AA;
 SQ

Query Match 100.0%; Score 32; DB 19; Length 74;
 Best Local Similarity 100.0%; Pred. No. 8.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTISS 7
 |||||
 Db 50 flltiss 56

RESULT 9
 AAR62923
 ID AAR62923 standard; protein; 75 AA.

XX AAR62923;

DT 18-JUL-1995 (first entry)

DE Human cytomegalovirus antibody light chain variable region.

KW Human cytomegalovirus; antibody light chain variable region;
 KM HMCV; Immunodassay; Immunotherapy.

OS Homo sapiens.

XX WO9425490-A.

XX 10-NOV-1994.

XX 29-APR-1994; 94WO-US04705.

XX 30-APR-1993; 93US-0055985.

XX (Scri) SCRIPPS RES INST.

PI Barbas C, Burton R, Burton DR, Williamson A, Burtoni R;

DR WPI; 1994-358194/44.

PT Human monoclonal antibodies (MAbs) against human cytomegalovirus
 PT also nucleic acids and cell lines producing the MAbs, useful in
 PT diagnosis and immunotherapy

XX Claim 6: Page 151, 171pp; English.

CC AAR62618-R62920 and AAR62921-R62923 are human cytomegalovirus (HMCV)
 CC antibody heavy chain and light chain variable regions respectively.
 CC A monoclonal antibody (Mab) containing a pair of these regions can be
 CC used to detect HCMV, and anti-human HCMV Abs in human patients via
 CC a competitive immunoassay. The MAbs may also be useful in
 CC immunotherapy.

XX Sequence 75 AA;
 SQ

Query Match 100.0%; Score 32; DB 15; Length 75;
 Best Local Similarity 100.0%; Pred. No. 8.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTISS 7
 |||||
 Db 37 flltiss 43

RESULT 10
 AAM80981
 ID AAM80981 standard; Protein; 76 AA.

XX AAM80981;

DT 30-MAR-1999 (first entry)

DE Variable kappa light region 012 encoded amino acid.

KM Human; epidermal growth factor receptor; tumour; EGF;
 KM transforming growth factor alpha; TGF-alpha.

OS Homo sapiens.

XX WO9850433-A2.

XX 12-NOV-1998.

XX 05-MAY-1998; 98WO-US09160.

XX 05-MAY-1997; 97US-0851362.

XX (ABGE-) ABGENIX INC.

PI Gallo M, Jakobovits A, Jia X, Yang X;

DR WPI; 1999-034712/03.

PT Humanised antibodies against epidermal growth factor receptor, EGF-r
 PT - useful to treat solid tumours whilst inducing reduced immunogenic
 PT or allergic effects compared to mouse or mouse-derived antibodies

XX Example 3; Page 105; 143pp; English.

CC The variable kappa light region 012 encoded amino acid was used in the
 CC production of anti-epidermal growth factor receptor (EGF-r)-antibodies.
 CC The antibodies can be administered therapeutically to patients (human or
 CC veterinary) to treat solid tumours. EGF-r is overexpressed on many human
 CC solid tumour types, and the fully human antibodies (i.e. comprising and
 CC inhibit both epidermal growth factor (EGF) and transforming growth factor
 CC alpha (TGF-alpha) binding to EGF-r (known to lead to cellular
 CC proliferation and tumour growth). They can prevent tumour cell growth
 CC and, in combination with an antineoplastic agent, may eradicate
 CC established tumours. The fully human antibodies can minimise the
 CC immunogenic and allergic responses intrinsic to previous mouse/rat or
 CC mouse/rat-derived antibodies.

XX Sequence 76 AA;
 SQ

Query Match 100.0%; Score 32; DB 20; Length 76;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLTISS 7
 |||||
 Db 52 flltiss 58

RESULT 11

AAM62807
 ID AAM62807 standard; Peptide; 82 AA.

XX AAM62807;

DT 23-SEP-1998 (first entry)

DE Amino acid sequence of a human antibody fragment.
 XX
 XX Human: Immunoglobulin Ig; transgenic; non-human mammal;
 KW Inactivated endogenous Ig locus; B-cell development;
 KW human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
 KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
 KW production; antibody.
 XX
 OS Homo sapiens.
 XX
 PN WO9824893-A2.
 PD 11-JUN-1998.
 XX
 PF 03-DEC-1997; 97WO-US23091.
 XX
 PR 03-DEC-1996; 96US-0759620.
 XX
 PA (ABGE-) ABGENIX INC.
 XX
 PI Green L, Jakobovits A, Klepholz S, Kucherlapati R;
 PI Mendez M;
 DR WPI; 1998-333314/29.
 XX
 PT New transgenic non-human mammals - having an inactivated
 PT immunoglobulin locus and a near complete human immunoglobulin locus,
 PT used for production of human antibodies
 XX
 PS Disclosure: Page 78; 128pp; English.
 XX
 CC AAM62793-822 represent fragments of human antibodies produced by
 CC transgenic Xenomice, created using the method of the invention. The
 CC specification describes a transgenic non-human mammal which has genome
 CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
 CC locus, so that the mammal does not display normal B-cell development. The
 CC modified genome also has an inserted human heavy chain Ig locus in
 CC germline configuration, the human heavy chain Ig locus comprising a human
 CC micro constant region and regulatory and switch sequences, human J-H
 CC genes, human D-H genes, and human V-H genes and an inserted human kappa
 CC light chain Ig locus in germline configuration, the human kappa light
 CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
 CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
 CC are selected to restore normal B-cell development in the mammal. The
 CC transgenic animals have a near complete human Ig locus, including both a
 CC human heavy chain locus and a human kappa light chain locus. They can
 CC be used for the production of human antibodies when exposed to
 CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
 CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
 CC respectively.
 CC
 SQ Sequence 82 AA;
 XX
 XX
 Query Match 100.0%; Score 32; DB 19; Length 82;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FTLTIS 7
 | | | | | | |
 Db 58 fcltiss 64
 XX
 RESULT 12
 AAM14491
 ID AAM14491 standard; Protein; 84 AA.
 XX
 AC AAM14491;
 XX
 DT 28-JAN-1997 (first entry)
 XX
 DE Monoclonal antibody D VK.
 XX

KW heavy; light chain; monoclonal antibody; antigen 3; marker; melanoma;
 KW permanent human tumour cell line; tumour-associated antigen; epitope;
 KW gastrointestinal tumour; pancreatic carcinoma; diagnostic; therapeutic;
 KW antigen 11; Vibrio cholera; neuraminidase-resistant; ganglioside GD2.
 XX
 OS Synthetic.
 XX
 PN EP727436-A1.
 PD 21-AUG-1996.
 XX
 PF 21-MAR-1990; 90EP-0105322.
 XX
 PR 24-MAR-1989; 89DE-3909799.
 XX
 PA (BEHM) BEHRINGWERKE AG.
 XX
 PI Auerbach B, Bosslet K, Sedlacek H, Seemann G;
 DR WPI; 1996-372836/38.
 DR N-PSDB; AAT63508.
 XX
 PT Monoclonal antibody to tumour-associated antigen - useful as
 PT gastrointestinal tumour marker
 XX
 PS Disclosure: Page 14; 19pp; German.
 XX
 CC AAM14490-91 are the heavy and light chains (respectively) of monoclonal
 CC antibody (Mab) D. Mab D recognises Vibrio cholera
 CC neuraminidase-resistant epitope of ganglioside GD2, from a human melanoma
 CC cell line. Mabs A, B and C (see AAM14484-89) are mentioned in the
 CC specification, but are not part of the claims. Mabs A and B recognise
 CC antigens 3 and 11 resp., of a permanent human tumour cell line. Mab C
 CC also recognises an epitope of a tumour-associated antigen. These antigens
 CC occur at high concns. in the serum of patients with gastrointestinal
 CC tumours, e.g. pancreatic carcinoma, and are thus useful as tumour markers
 CC for diagnostic or therapeutic purposes.
 CC
 SQ Sequence 84 AA;
 XX
 XX
 Query Match 100.0%; Score 32; DB 17; Length 84;
 Best Local Similarity 100.0%; Pred. No. 9.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FTLTIS 7
 | | | | | | |
 Db 49 fcltiss 55
 XX
 RESULT 13
 AAR9878
 ID AAR9878 standard; Protein; 84 AA.
 XX
 AC AAR9878;
 XX
 DT 28-JAN-1997 (first entry)
 XX
 DE Monoclonal antibody D VK.
 XX
 KW Monoclonal antibody; Mab; epitope; tumour-associated antigen;
 KW marker; antigen.
 XX
 OS Synthetic.
 XX
 PN EP727435-A1.
 PD 21-AUG-1996.
 XX
 PF 21-MAR-1990; 90EP-0105322.
 XX
 PR 24-MAR-1989; 89DE-3909799.
 XX

PA (BEHW) BEHRINGER AG.
XX
XX Auerbach B, Bosslet K, Sedlacek H, Seemann G;
XX
XX WPI; 1996-372835/38.
DR N-PSDB; AAT36666.
XX
XX Monoclonal antibody to tumour-associated antigen - useful as
PT gastrointestinal tumour marker
XX
XX
XX Disclosure; Page 14; 19pp; German.
XX
XX Mab C (AAT36659-T36660) is a monoclonal antibody that recognises an
CC epitope of a tumour-associated antigen occurring at high concn. in
CC the serum of patients with gastrointestinal tumours, e.g. pancreatic
CC carcinoma, and is thus useful as a tumour marker for diagnostic or
CC therapeutic purposes.
CC
CC Mabs A, B and D are mentioned in the specification, but are not
CC part of the claims.
CC
CC Mab A (AAT36661-T36662) recognises antigen 3 of permanent human
CC tumour cell line.
CC
CC Mab B (AAT36663-T36664) recognises antigen 11 of permanent human
CC tumour cell line.
CC
CC Mab D (AAT36665-T36666) recognises a Vibrio cholera neuraminidase-
CC resistant epitope of ganglioside GD2, from a human melanoma cell
CC line.
XX
XX
SQ Sequence 84 AA:

Query Match 100.0%; Score 32; DB 17; Length 84;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FTLTSS 7
DB 49 flltss 55

RESULT 14
AAM62806
ID AAM62806 standard; Peptide: 86 AA.
XX
XX AAM62806;
XX
XX 23-SEP-1998 (first entry)
XX
XX
XX Amino acid sequence of a human antibody fragment.
DE
XX
XX Human; immunoglobulin; Ig; transgenic; non-human mammal;
XX
XX Inactivated endogenous Ig locus; B-cell development;
XX
XX human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
KW production; antibody.
XX
XX
XX Homo sapiens.
OS
XX
XX WO9824893-A2.
PN
XX
XX 11-JUN-1998.
PD
XX
XX 03-DEC-1997; 97WO-US23091.
PF
XX
XX 03-DEC-1996; 96US-0759620.
PR
XX
XX (ABGE-) ABGENIX INC.
PA
XX
XX Green L, Jakobovits A, Klapholz S, Kucherlapati R;
PI Mendez M;
PI WPI; 1998-333314/29.
XX
XX
XX New transgenic non-human mammals - having an inactivated

PT immunoglobulin locus and a near complete human immunoglobulin locus,
PT used for production of human antibodies
XX
XX
XX Disclosure; Page 78; 128pp; English.
XX
XX AAM62793-822 represent fragments of human antibodies produced by
CC transgenic Xenomice, created using the method of the invention. The
CC specification describes a transgenic non-human mammal which has genome
CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
CC locus, so that the mammal does not display normal B-cell development. The
CC modified genome also has an inserted human heavy chain Ig locus in
CC germ-line configuration, the human heavy chain Ig locus comprising a
CC micro constant region and regulatory and switch sequences, human J-H
CC genes, human D-H genes, and human V-H genes and an inserted human kappa
CC light chain Ig locus in germ-line configuration, the human kappa light
CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
CC are selected to restore normal B-cell development in the mammal. The
CC transgenic animals have a near complete human Ig locus, including both a
CC human heavy chain locus and a human kappa light chain locus. They can
CC be used for the production of human antibodies when exposed to
CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
CC respectively.
XX
XX
SQ Sequence 86 AA:

Query Match 100.0%; Score 32; DB 19; Length 86;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FTLTSS 7
DB 50 flltss 56

RESULT 15
AAM59619
ID AAM59619 standard; Protein: 88 AA.
XX
XX AAM59619;
XX
XX 12-OCT-1998 (first entry)
XX
XX
XX Anti-RSV F protein Hu19 light chain GL DPK9.
DE
XX
XX Monoclonal antibody; human; Hu19; engineered antibody; RSV;
KW respiratory syncytial virus; fusion protein; infection;
KW complementarity determining region; CDR; therapy; diagnosis.
XX
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Region 24..34
FT /label= CDR1
FT 50..57
FT /label= CDR2
XX
XX WO9819704-A1.
PN
XX
XX 14-MAY-1998.
PD
XX
XX 23-OCT-1997; 97WO-US19203.
PF
XX
XX 01-NOV-1996; 96US-0030149.
PR
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX
XX Deen KC, Dillon SB, Porter TG, Sweet RW;
PI WPI; 1998-286600/25.
XX
XX
XX New transgenic non-human mammals - having an inactivated

PT Monoclonal antibodies reactive with Respiratory Syncytial Virus -
 PT useful for detection, prevention and treatment of RSV infections
 PS
 PS Disclosure: Fig 3; 109pp; English.
 XX

CC This is the heavy chain amino acid sequence of germline (GL)
 CC DPK9 antibody. The invention relates to the construction and use
 CC of human Mabs specific for the fusion (F) protein of respiratory
 CC syncytial virus (RSV) to passively treat, prevent or detect RSV
 CC infection. Hu19A, Hu19B, Hu19C and Hu19D Mabs are claimed. These
 CC are reshaped human antibodies comprising a heavy chain selected
 CC from 19A, 19B, 19C or 19D (see AAW59615-18), which are based on the
 CC GL Dp58 sequence, and a light chain selected from 19A, 19B, 19C or
 CC 19D (see AAW59620-21). Such engineered antibodies are neutralising;
 CC they inhibit virus growth in vitro and in vivo in animal models of
 CC RSV infection. Nucleic acids encoding the human Mabs, recombinant
 CC plasmids (see AAW41427-33) and host cells (e.g. COS, CHO, myeloma)
 CC are provided.
 CC
 XX

SQ Sequence 88 AA;

Query Match 100.0%; Score 32; DB 19; Length 88;
 Best local Similarity 100.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTUTISS 7
 |||||
 Db 71 ftltiss 77

Search completed: July 15, 2002, 12:57:54
 Job time: 410 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:00:57 ; Search time 95.45 Seconds
(without alignments)
7.047 Million cell updates/sec

Title: US-09-712-819a-2

Perfect score: 33

Sequence: 1 TRIFSKL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	97.0	529	2 T03879	hypothetical prote
2	30	90.9	499	2 T03877	hypothetical prote
3	30	90.9	1106	2 C87128	probable transcrip
4	30	90.9	1227	2 T50394	DNA-directed RNA p
5	29	87.9	244	2 B87417	hypothetical prote
6	29	87.9	291	2 D96998	phosphatidylserine
7	29	87.9	352	2 T32314	hypothetical prote
8	29	87.9	590	2 E90089	hypothetical prote
9	29	87.9	592	2 S42220	transcription init
10	29	87.9	1071	2 T18597	hypothetical prote
11	28	84.8	39	2 H81081	hypothetical prote
12	28	84.8	179	2 B75636	transcription regu
13	28	84.8	187	2 G70196	hypothetical prote
14	28	84.8	193	2 G84869	hypothetical prote
15	28	84.8	328	2 E88930	hypothetical prote
16	28	84.8	381	2 T39187	neuropeptide Y/pep
17	28	84.8	422	2 T21820	hypothetical prote
18	28	84.8	483	2 S25070	xylokinase (EC 2
19	28	84.8	586	2 T04716	hypothetical prote
20	28	84.8	744	2 T00652	hypothetical prote
21	28	84.8	834	2 T00512	serine/threonine-s
22	28	84.8	852	2 AD0050	probable phosphoen
23	28	84.8	1317	2 A54831	nuclear pore compl
24	28	81.8	156	2 C69330	conserved hypotet
25	27	81.8	204	2 D71648	hypothetical prote
26	27	81.8	237	2 S73253	hypothetical prote
27	27	81.8	342	2 T88045	hypothetical prote
28	27	81.8	343	2 A41748	limicin precursor
29	27	81.8	352	2 S70972	bipe protein - Esc

30	27	81.8	355	2 T10722	anthocyanidin synt
31	27	81.8	392	2 C96766	hypothetical prote
32	27	81.8	398	2 F85056	hypothetical prote
33	27	81.8	405	2 D96709	probable B-box zin
34	27	81.8	416	2 E86384	probable zinc fing
35	27	81.8	440	2 AF2934	uracil transport p
36	27	81.8	463	2 S52754	aromatic amino aci
37	27	81.8	502	2 A98348	glucosylceramidase
38	27	81.8	522	2 T18583	prolyl-tRNA synth
39	27	81.8	582	2 B82805	complement C8 alph
40	27	81.8	584	1 C8H0A	hypothetical prote
41	27	81.8	611	2 A83926	protein F7F22.2 [i
42	27	81.8	612	2 H96494	hypothetical prote
43	26	78.8	67	2 AB2780	ORF MSY131 hypote
44	26	78.8	85	2 T28292	hypothetical prote
45	26	78.8	121	2 T50205	

ALIGNMENTS

RESULT 1

T03879
hypothetical protein F07G11.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #extl_change 26-May-2000

C:Accession: T03879

R:Sammons, L.; Wohldmann, P.; Sansone, J.

Submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid F07G11.

A:Reference number: Z15127

A:Accession: T03879

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-529 <SAM>

A:Cross-references: EMBL:AF016419; NID:g2291159; PIDN:AA65295.1; PID:g2291160

C:Genetics:

A:Map position: V

A:Insertions: 73/2; 120/2; 155/3; 193/2; 219/3; 269/1; 392/3

A:Note: F07G11.3
C:Superfamily: Caenorhabditis elegans hypothetical protein F07G11.3

Query Match 97.0%; Score 32; DB 2; Length 529;
Best Local Similarity 85.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSKL 7
DB 440 TRIFSKL 446

RESULT 2

T03877
hypothetical protein F07G11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #extl_change 26-May-2000

C:Accession: T03877

R:Sammons, L.; Wohldmann, P.; Sansone, J.

Submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid F07G11.

A:Reference number: Z15127

A:Accession: T03877

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-499 <SAM>

A:Cross-references: EMBL:AF016419; NID:g2291159; PIDN:AA65298.1; PID:g2291163

C:Genetics:

A:Map position: V

A:Insertions: 56/2; 103/2; 138/3; 176/2; 202/3; 246/3; 318/3; 363/3

A:Note: F07G11.4
C:Superfamily: Caenorhabditis elegans hypothetical protein F07G11.3

Query Match 90.9% Score 30; DB 2; Length 499;
Best Local Similarity 85.7% Pred. No. 45;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRIFSRL 7
||:||||
Db 410 TRIFSRL 416

RESULT 3

Probable transcription regulator [Imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87128
R:COLE, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duchoy, S.; Feltewell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C87128
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1106 <STO>
A:Cross-references: GB:AL450380; NID:g13093492; PIDN:GAC30706.1; GSPDB:GN00147
C:Genetics:
A:Gene: ML1753

Query Match 90.9% Score 30; DB 2; Length 1106;
Best Local Similarity 85.7% Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRIFSRL 7
||:||||
Db 1084 TRIFSRL 1090

RESULT 4

DNA-directed RNA polymerase subunit [Imported] - fission yeast (Schizosaccharomyces pombe
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: T50394
R:Beck, A.; Borzym, K.; Reinhardt, R.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z25067
A:Accession: T50394
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1227 <BEC>
A:Cross-references: EMBL:AL136535; PIDN:CAB66435.1; GSPDB:GN00067; SPDB:SPBP23A10.07
C:Genetics:
A:Gene: rpa2; SPDB:SPBP23A10.07
A:Map position: 2
C:Superfamily: DNA-directed RNA polymerase 132k polypeptide

Query Match 90.9% Score 30; DB 2; Length 1227;
Best Local Similarity 71.4% Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRIFSRL 7
||:||||
Db 476 TRIFSRL 482

RESULT 5

887417

hypothetical protein CC1353 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: B87417

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87417

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-244 <STO>
A:Cross-references: GB:AE005673; NID:g13422702; PIDN:AAK23334.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1353

Query Match 87.9% Score 29; DB 2; Length 244;
Best Local Similarity 100.0% Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRFSK 6
|||||
Db 69 TRFSK 74

RESULT 6

Phosphatidylserine decarboxylase [Imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D96998
R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4836, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D96998
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-291 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78775.1; PID:g15023687; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0799
C:Superfamily: Escherichia coli phosphatidylserine decarboxylase

Query Match 87.9% Score 29; DB 2; Length 291;
Best Local Similarity 71.4% Pred. No. 45;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TRFSKL 7
||:||||
Db 43 TRFSKL 49

RESULT 7

hypothetical protein F31f4.14 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T32314
R:Blanchard, M.; Kramer, J.; Elliott, G.; Teyman, B.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F31f4.
A:Reference number: Z21149
A:Accession: T32314
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-352 <BLA>

A:Cross-references: EMBL:AF024503; PIDN:AA070390.1; GSPDB:GN00023; CESP:F31F4.14
 A:Experimental source: strain Bristol N2; clone F31F4
 C:Genetics:
 A:Gene: CESP:F31F4.14
 A:Map position: 5
 A:Introns: 100/3; 170/3; 223/3; 293/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein C33A12.9b

Query Match 87.9%; Score 29; DB 2; Length 352;
 Best Local Similarity 85.7%; Pred. No. 55;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 :|||||
 DB 48 SRIFSKL 54

RESULT 8
 E90089
 hypothetical protein orf590 [imported] - Guillardia theta nucleomorph
 C:Species: nucleomorph guillardia theta
 A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
 C:Accession: E90089
 R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
 Nature 410, 1091-1096, 2001
 A:Title: The highly reduced genome of an enslaved algal nucleus.
 A:Reference number: A99082; MUID:11323671
 A:Accession: E90089
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-590 <DDU>
 A:Cross-references: GB:AF165818; NID:913794473; PIDN:AAK39848.1; GSPDB:GN00150
 C:Genetics:
 A:Gene: orf590
 A:Map position: 1
 A:Genome: nucleomorph
 C:Keywords: nucleomorph

Query Match 87.9%; Score 29; DB 2; Length 590;
 Best Local Similarity 85.7%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 :|||||
 DB 203 SRIFSKL 209

RESULT 9
 S42220
 transcription initiation factor IID chain p62 - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 20-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 17-Nov-2000
 C:Accession: S42220
 R:Kokubo, T.; Gong, D.W.; Wootton, J.C.; Horikoshi, M.; Roeder, R.G.; Nakatani, Y.
 Nature 367, 484-487, 1994
 A:Title: Molecular cloning of Drosophila TFIID subunits.
 A:Reference number: S42220; MUID:94150630
 A:Accession: S42220
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-592 <KOK>
 A:Cross-references: EMBL:U06459; NID:9458681; PIDN:AA046480.1; PTD:9458682
 C:Genetics:
 A:Gene: FlyBase:Taf40
 A:Cross-references: FlyBase:FBgn0000617
 C:Keywords: transcription initiation

Query Match 87.9%; Score 29; DB 2; Length 592;
 Best Local Similarity 100.0%; Pred. No. 92;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TRIFSK 6
 :|||||
 DB 337 TRIFSK 342

RESULT 10
 T18597
 hypothetical protein AC3.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18597
 R:McMurray, A.
 submitted to the EMBL Data Library, April 1996
 A:Reference number: Z18995
 A:Accession: T18597
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1071 <MTL>
 A:Cross-references: EMBL:Z71177; PIDN:CAA94872.1; GSPDB:GN00023; CESP:AC3.5
 A:Experimental source: clone AC3
 C:Genetics:
 A:Gene: CESP:AC3.5
 A:Map position: 5
 A:Introns: 19/3; 124/2; 164/2; 222/3; 376/2; 408/3; 449/1; 537/3; 604/3; 700/2; 808/3

Query Match 87.9%; Score 29; DB 2; Length 1071;
 Best Local Similarity 85.7%; Pred. No. 1,7402;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 :|||||
 DB 788 TRIFSKL 794

RESULT 11
 H81081
 hypothetical protein NMB1455 [imported] - Neisseria meningitidis (strain MC58 serogro
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: H81081
 R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
 Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
 ri, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizze, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755
 A:Accession: H81081
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-39 <TEF>
 A:Cross-references: GB:AE002495; GB:AE002098; NID:97226690; PIDN:AAF41814.1; PTD:9722
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1455

Query Match 84.8%; Score 28; DB 2; Length 39;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIFSKL 7
 :|||||
 DB 24 RIFSKL 29

RESULT 12
 B75636
 transcription regulator, Gere family - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75636
 R:White, O.; Eelsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Mc
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* RL.
 A:Reference number: A75250; MUID:20036896
 A:Accession: B75636
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-179 <NH>
 A:Cross-references: GB:AE001827; NID:96460959; PIDN:AAF12666.1; PID:96460963; TIGR:DRC00
 C:Genetics:
 A:Experimental source: strain RL
 C:Genetics:
 A:Gene: DRC0012
 A:Map position: plasmid
 A:Genome: plasmid
 A:Note: plasmid CPl

Query Match 84.8%; Score 28; DB 2; Length 179;
 Best Local Similarity 71.4%; Pred. No. 48;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 |||
 Db 150 TRVFOKL 156

RESULT 13
 G70196
 hypothetical protein BB0776 - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: G70196
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Klevavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.
 A:Reference number: A70100; MUID:9605943
 A:Accession: G70196
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-187 <KLE>
 A:Cross-references: GB:AE001177; GB:AE000783; NID:g2688711; PIDN:AA67137.1; PID:g268872
 A:Experimental source: strain B31

Query Match 84.8%; Score 28; DB 2; Length 187;
 Best Local Similarity 83.3%; Pred. No. 50;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6
 |||
 Db 3 TRVFSK 8

RESULT 14
 G84869
 hypothetical protein At2g43730 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84869
 R:Liu, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 Euse, D.; Nierman, W.C.; White, O.; Eelsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: AB4420; MUID:20083487
 A:Accession: G84869

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-193 <STO>
 A:Cross-references: GB:AE002093; NID:g2281097; PIDN:AA64033.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g43730
 A:Map position: 2

Query Match 84.8%; Score 28; DB 2; Length 193;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RIFSKL 7
 |||||
 Db 173 RIFSKL 178

RESULT 15
 E88930
 protein R1G11.11 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: E88930
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C-
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A:Accession: E88930
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-328 <STO>
 A:Cross-references: GB:chr_V; PIDN:AA69085.1; PID:g2384860; GSPDB:GN00023; CESP:R11G
 C:Genetics:
 A:Gene: R1G11.11
 A:Map position: 5

Query Match 84.8%; Score 28; DB 2; Length 328;
 Best Local Similarity 85.7%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 |||||
 Db 280 TRIFSKL 286

Search completed: July 15, 2002, 13:00:59
 Job time: 465 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:23:30 ; Search time 44.9 Seconds

(without alignments)
6.036 Million cell updates/sec

Title: US-09-712-819A-2

Sequence: 1 TRIFSKL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	93.9	141	1	HBB_DASAK
2	30	90.9	1227	1	RPA2_SCHPO
3	29	87.9	592	1	T2D5_DROME
4	28	84.8	343	1	ILVC_BUCML
5	28	84.8	381	1	NY2R_HUMAN
6	28	84.8	381	1	NY2R_MACMU
7	28	84.8	384	1	NY2R_BOVIN
8	28	84.8	385	1	NY2R_MOUSE
9	28	84.8	438	1	IAP1_DROME
10	28	84.8	483	1	XYLB_KLEPN
11	28	84.8	476	1	SCPI_HUMAN
12	28	84.8	1317	1	N145_YEAST
13	27	81.8	184	1	CLIP_MOUSE
14	27	81.8	204	1	Y860_RICPR
15	27	81.8	237	1	YC28_PORPU
16	27	81.8	338	1	LDM_HUMAN
17	27	81.8	343	1	LDM_CHICK
18	27	81.8	463	1	AROP_CORGL
19	27	81.8	584	1	COBA_HUMAN
20	26	78.8	203	1	YEVB_YEAST
21	26	78.8	208	1	YC29_GUITH
22	26	78.8	217	1	GTML_RAT
23	26	78.8	369	1	HIS8_ZYMO
24	26	78.8	389	1	CD34_CANFA
25	26	78.8	396	1	PSD4_DROME
26	26	78.8	413	1	HP55_TAMST
27	26	78.8	449	1	SR54_RICPR
28	26	78.8	459	1	TRPB_PYRHO
29	26	78.8	476	1	CHK1_HUMAN
30	26	78.8	476	1	CHK1_MOUSE
31	26	78.8	479	1	ACH9_HUMAN
32	26	78.8	479	1	ACH9_RAT
33	26	78.8	580	1	P69_MYCHR

34	26	78.8	607	1	UVRC_PSEFL	P32966 pseudomonas
35	26	78.8	622	1	VEIN_DROME	Q94918 drosophila
36	26	78.8	638	1	YDEC_SCHPO	Q10445 schizosacch
37	26	78.8	711	1	FRE2_YEAST	P36033 saccharomyc
38	26	78.8	891	1	DPO1_HELPY	P56105 helicobacte
39	26	78.8	897	1	DPO1_HELPY	Q92369 helicobacte
40	26	78.8	946	1	YVIA2_YEAST	P40559 saccharomyc
41	26	78.8	2731	1	RPA2_CVMJH	P29982 murine coro
42	26	78.8	3358	1	PGCV_MOUSE	Q62059 mus musculu
43	26	78.8	4485	1	DYHG_CHIRE	Q39575 chlamydomon
44	25	75.8	82	1	VC14_VACCC	P21045 vaccinia vi
45	25	75.8	122	1	PQRA_PROVU	Q52620 proteus vul

ALIGNMENTS

RESULT	ID	BBB_DASAK	STANDARD	PRT	141 AA
AC	P56692:				
DT	15-JUL-1999 (Rel. 38, Last Sequence update)				
DT	15-JUL-1999 (Rel. 38, Last Sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Hemoglobin beta chain.				
GN	HBB.				
OS	Dasyatis akajei (Red stingray) (Akai).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;				
OC	Elasmobranchii; Squalae; Hyposquales; Pristiogaster; Batoidae;				
OC	Myliobatiformes; Myliobatoidei; Dasyatidae; Dasyatis.				
OX	NCBI_TaxID=31902;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).				
RC	TISSUE=Blood;				
RA	Chong K.T., Miyazaki G., Morimoto H., Oda Y., Park S., Tsukihara T.;				
RT	"The three dimensional structures of deoxy and CO forms of hemoglobin				
RL	from Dasyatis akajei, a cartilaginous fish."				
CC	Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.				
CC	- FUNCTION: INVOLVED IN OXYGEN TRANSPORT FROM THE LUNG TO THE				
CC	VARIOUS PERIPHERAL TISSUES.				
CC	- SUBUNIT: HETEROQUATERNARY OF TWO ALPHA CHAINS AND TWO BETA CHAINS.				
CC	- TISSUE SPECIFICITY: RED BLOOD CELLS.				
CC	- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: AB023723; BAA75250.1; -				
DR	PDB: 1CG5; 01-APR-99.				
DR	PDB: 1CG8; 01-APR-99.				
DR	InterPro: IPR002337; Beta_haem.				
DR	InterPro: IPR000971; Globin.				
DR	Pfam: PF00042; globin; 1.				
DR	PRINTS: PR00814; BETAHAEM.				
DR	PROSITE: PS01033; GLOBIN; 1.				
KW	Heme; Oxygen transport; Transport; Erythrocyte;				
KW	3D-structure.				
FT	INIT_MET 0				
FT	METAL 59 59 IRON (HEME DISTAL LIGAND).				
FT	METAL 88 88 IRON (HEME PROXIMAL LIGAND).				
SQ	SEQUENCE 141 AA; 16289 MW; DAED4578804027B CRC64;				
QY	1 TRIFSKL 7				

Query Match 93.9%; Score 31; DB 1; Length 141;

Best local Similarity 85.7%; Pred. No. 4.2;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Db          39 TRUESKL 45

RESULT      2
RPA2_SCHPO  STANDARD:      PRT: 1227 AA.
ID          16-OCT-2001 (Rel. 40, Created)
AC          09p7x8;
DE          16-OCT-2001 (Rel. 40, Last sequence update)
DT          16-OCT-2001 (Rel. 40, Last annotation update)
DE          Probable DNA-directed RNA polymerase I polypeptide 2 (EC 2.7.7.6) (RNA
GN          polymerase I subunit 2).
RPA2 OR SPBP23A10.07.
OS          Schizosaccharomyces pombe (Fission yeast).
OC          Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC          Schizosaccharomycetales; Schizosaccharomycetaceae;
OC          Schizosaccharomycetes.
OX          NCBI_TaxID=4896;
RN          (1)
RP          SEQUENCE FROM N.A.
RC          STRAIN=972;
RA          Beck A., Borzym K., Reinhardt R., McDougall R.C., Rajandream M.A.,
RA          Barrell B.G.;
RL          Submitted (Jan-1999) to the EMBL/Genbank/DBJ databases.
CC          -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC          OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC          SUBSTRATES. RNA POLYMERASE I IS ESSENTIALLY USED TO TRANSCRIBE
CC          RIBOSOMAL DNA UNITS.
CC          -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC          (RNA)(n).
CC          -1- SUBUNIT: RNA POLYMERASE I CONSISTS OF 14 DIFFERENT SUBUNITS. THIS
CC          SUBUNIT IS THE SECOND LARGEST COMPONENT OF RNA POLYMERASE I (BY
CC          SIMILARITY).
CC          -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC          -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE
CC          FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC          PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC          III FOR 5S AND TRNA GENES.
CC          -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
CC          -----
CC          This SWISS-PROT entry is copyright. It is produced through a collaboration
CC          between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC          the European Bioinformatics Institute. There are no restrictions on its
CC          use by non-profit institutions as long as its content is in no way
CC          modified and this statement is not removed. Usage by and for commercial
CC          entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC          or send an email to license@isb-sib.ch).
CC          -----
DR          EMBL: AL136535; CAB66435.1.
DR          InterPro: IPR001572; RNA_POL_B.
DR          Pfam: PF00562; RNA_POL_B; 1.
DR          PROSITE: PS01166; RNA_POL_BETA; 1.
KM          Transferrase: DNA-directed RNA polymerase: Transcription; zinc;
KW          Zinc-finger; Metal-binding; C4-type (POTENTIAL).
FT          ZN_Fing 1142 1171
FT          SEQUENCE 1227 AA; 137739 MW; 614A45ADE74D3419 CRC64;
SO          Query Match          90.9%; Score 30; DB 1; Length 1227;
SO          Best Local Similarity 71.4%; Pred. No. 54;
SO          Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

DT          01-OCT-1996 (Rel. 34, Last sequence update)
DT          16-OCT-2001 (Rel. 40, Last annotation update)
DE          Transcription initiation factor TFIID 62 kDa subunit (P62) (TAFII-60).
DE          TAF60 OR CG9348.
GN          Drosophila melanogaster (Fruit fly).
OS          Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC          Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC          Ephydroidea; Drosophilidae; Drosophila.
OX          NCBI_TaxID=7227;
RN          (1)
RP          SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC          MEDLINE=94150630; PubMed=7545910;
RA          Kokubo T., Gong D.-W., Woolton J.C., Horikoshi M., Roeder R.G.,
RA          Nakatani Y.;
RT          "Molecular cloning of Drosophila TFIID subunits.";
RL          Nature 367:484-487(1994).
RN          (2)
RP          SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC          STRAIN=OREGON-R;
RC          MEDLINE=94085406; PubMed=8262073;
RA          Weinzierl R.O., Ruppert S., Dynlacht B.D., Tanese N., Tian R.;
RT          "Cloning and expression of drosophila TAFII60 and human TAFII70
RT          reveal conserved interactions with other subunits of TFIID.";
RL          EMBO J. 12:5303-5309(1993).
RN          (3)
RP          SEQUENCE FROM N.A.
RC          STRAIN=BERKELEY;
RC          MEDLINE=20196006; PubMed=10731132;
RA          Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA          Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA          George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA          Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA          Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA          Abail J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
RA          Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA          Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA          Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA          Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA          Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA          de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA          Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA          Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA          Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA          Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA          Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA          Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA          Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA          Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA          Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA          Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA          Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA          Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA          Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA          Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA          Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA          Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA          Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA          Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA          Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA          Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA          Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA          Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA          Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT          "The genome sequence of Drosophila melanogaster.";
RL          Science 287:2185-2195(2000).
CC          -1- FUNCTION: TFIID IS A MULTIMERIC PROTEIN COMPLEX THAT PLAYS A
CC          CENTRAL ROLE IN MEDIATING PROMOTER RESPONSES TO VARIOUS ACTIVATORS
CC          AND REPRESSORS. BINDS TIGHTLY TO TAFII-250 AND ALSO DIRECTLY
CC          INTERACTS WITH TAFII-40.
CC          -1- SUBUNIT: TFIID IS COMPOSED OF TATA BINDING PROTEIN (TBP) AND A
CC          NUMBER OF TBP-ASSOCIATED FACTORS (TAFs).
CC          -1- SUBCELLULAR LOCATION: Nuclear.

```

CC -1- SIMILARITY: BELONGS TO THE TAF2E FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U06459; AAC6480.1; -.
DR EMBL: L25443; AAA16536.1; -.
DR EMBL: AE003516; AAF49139.1; -.
DR TRANSFAC: T02124; -.
DR FLYBASE: FBgn0010417; Taf60.
DR InterPro: IPR00166; Histone_core.
DR Pfam: PF02369; TAF; 1.
KW Transcription regulation; Nuclear protein.
FT CONFLICT 87 87 F -> L (IN REF. 2).
FT CONFLICT 568 568 E -> D (IN REF. 3).
SQ SEQUENCE 592 AA; 64322 MW; 5DBF07AE35E0F7E3 CRC64;

Query Match 87.9%; Score 29; DB 1; Length 592;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRFSK 6
DB 337 TRFSK 342

RESULT 4
ID ILVC_BUCML STANDARD; PRT; 343 AA.
AC Q9A096;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Ketol-acid reductoisomerase (EC 1.1.1.86) (Acetohydroxy-acid
DE isomeroeductase) (Alpha-keto-beta-hydroxyacid reductoisomerase)
DE (Fragment).
GN ILVC.
OS Buchnera aphidicola (subsp. Macrosiphoniella ludovicianaee).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118105;
RN [1]
RP SEQUENCE FROM N.A.
RA Wernegreen J.J., Moran N.A.;
RT "Accelerated evolutionary rates at biosynthetic loci of Buchnera-
RT uroleucon";
RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)
CC -1- (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS; SECOND STEP.
CC -1- SIMILARITY: BELONGS TO THE KETOL-ACID REDUCTOISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF21557; AAK01028.1; -.
DR InterPro: IPR000506; Acetohydrox-ac_isomrctse.
DR Pfam: PF01450; ILVC; 1.
KW Oxidoreductase; Branched-chain amino acid biosynthesis; NADP.
FT NON_TER 1 1
FT ACT_SITE 65 65 POTENTIAL.
FT NON_TER 343 343
SQ SEQUENCE 343 AA; 39012 MW; 67369DC0E240113C CRC64;

Query Match 84.8%; Score 28; DB 1; Length 343;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RIFSKL 7
DB 227 RIFSKL 232

RESULT 5
ID NY2R_HUMAN STANDARD; PRT; 381 AA.
AC P49146; O13281; O13457; 09UE67;
DT 01-FEB-1996 (Rel. 33; Created)
DT 01-FEB-1996 (Rel. 33; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=96070760; PubMed=7592910;
RA Gerald C., Walker M.W., Vaysse P.J.-J., He C., Branchek T.A.,
RA Weisshank R.L.;
RT "Expression cloning and pharmacological characterization of a human
RT hippocampal neuropeptide Y/peptide YY Y2 receptor subtype";
RL J. Biol. Chem. 270:26758-26761(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96226058; PubMed=8632753;
RA Gehlert D.R., Beavers L.S., Johnson D., Gackenhelmer S.L.,
RA Schober D.A., Galski R.A.;
RT "Expression cloning of a human brain neuropeptide Y Y2 receptor";
RL Mol. Pharmacol. 49:224-228(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96032678; PubMed=7559383;
RA Rose P.M., Fernandes P., Lynch J.S., Frazier S.T., Fisher S.M.,
RA Kodukula K., Kienzie B., Seethala R.;
RT "Cloning and functional expression of a cDNA encoding a human type 2
RT neuropeptide Y receptor";
RL J. Biol. Chem. 270:22661-22664(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96209788; PubMed=8643460;
RA Yan H., Yang J., Marasco J., Yamaguchi K., Brenner S., Collins F.,
RA Katson W.;
RT "Cloning and functional expression of cDNAs encoding human and rat
RT pancreatic polypeptide receptors";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4661-4665(1996).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=97131518; PubMed=8975716;
RA Ammar D.A., Eadie D.M., Wong D.J., Ma Y.-Y., Kolakowski L.F. Jr.,
RA Yang-Feng T.L., Thompson D.A.;
RT "Characterization of the human type 2 neuropeptide Y receptor gene
RT (NPY2R) and localization to the chromosome 4q region containing the
RT type 1 neuropeptide Y receptor gene";
RL Genomics 38:392-398(1996).
RN [6]
RP SEQUENCE FROM N.A.
RA Zaslavsky R.L.;
RT "Human neuropeptide Y Y2 receptor gene";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE RANK
CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS

CC pty > NPY > PYY (3-36) > NPY (2-36) > [ILE-31, GLN-34] PP >
CC [ILEU-31, PRO-34] NPY > PP, [PRO-34] PYY AND NPY FREE ACID.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: HIGH LEVELS IN AMYGDALA, CORPUS CALLOSUM,
CC HIPPOCAMPUS AND SUBTHALAMIC NUCLEUS. ALSO DETECTABLE IN CAUDATE
CC NUCLEUS, HYPOTHALAMUS AND SUBSTANTIA NIGRA.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U36289; AAC50281.1; -
CC EMBL: U42766; AAB04120.1; -
CC EMBL: U42500; AAA93170.1; -
CC EMBL: U42389; AAB07760.1; -
CC EMBL: U50146; AAC51115.1; -
CC EMBL: U76254; AAD00248.1; -
CC GCRDB: GCR_1945; -
CC GCRDB: GCR_1946; -
CC GCRDB: GCR_1958; -
CC GCRDB: GCR_1974; -
CC GCRDB: GCR_2066; -
CC MIM: 162642; -
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1; 1.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PROSITE: PS00237; G-PROTEIN_RECPR_F1_1; 1.
CC PROSITE: PS50262; G-PROTEIN_RECPR_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC FT DOMAIN 1 49
CC TRANSSEM 73 82
CC FT DOMAIN 50 72
CC TRANSSEM 83 104
CC FT DOMAIN 105 124
CC TRANSSEM 125 146
CC FT DOMAIN 147 166
CC TRANSSEM 167 187
CC FT DOMAIN 188 214
CC TRANSSEM 215 240
CC FT DOMAIN 241 268
CC TRANSSEM 269 291
CC FT DOMAIN 292 304
CC TRANSSEM 305 328
CC FT DOMAIN 329 381
CC CARBOHYD 11 11
CC DISULFID 123 203
CC FT LIPID 342 342
CC CONFLICT 134 134
CC CONFLICT 172 172
CC CONFLICT 174 174
CC CONFLICT 202 202
CC CONFLICT 209 209
CC CONFLICT 248 248
CC CONFLICT 311 311
CC SEQUENCE 381 AA; 42731 MW; 7D018C0169597BC7 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 381;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
DB 240 TRIMSKL 246

RESULT 6
ID N2R_MACMU
NT2R_MACMU STANDARD: PRT: 381 AA.
AC Q96K74; 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21184974; PubMed-11287088;
RA Gehlert D.R., Yang P., George C., Mang Y., Schober D.,
RA Gackenhauer S., Johnson D., Beavers L.S., Gadsdill R.A., Baez M.;
RT "Cloning and characterization of Rhesus monkey neuropeptide Y receptor
RT subtypes";
RL Peptides 22:343-350(2001).
CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF303090; AAG40772.1; -
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm_1; 1.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PROSITE: PS00237; G-PROTEIN_RECPR_F1_1; 1.
CC PROSITE: PS50262; G-PROTEIN_RECPR_F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC FT DOMAIN 1 49
CC TRANSSEM 50 72
CC FT DOMAIN 73 82
CC TRANSSEM 83 104
CC FT DOMAIN 105 124
CC TRANSSEM 125 146
CC FT DOMAIN 147 166
CC TRANSSEM 167 187
CC FT DOMAIN 188 214
CC TRANSSEM 215 240
CC FT DOMAIN 241 268
CC TRANSSEM 269 291
CC FT DOMAIN 292 304
CC TRANSSEM 305 328
CC FT DOMAIN 329 381
CC CARBOHYD 11 11
CC DISULFID 123 203
CC FT LIPID 342 342
CC SEQUENCE 381 AA; 42729 MW; D689D691DC05A9D CRC64;

Query Match 84.8%; Score 28; DB 1; Length 381;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
DB 240 TRIMSKL 246


```

RESULT 7
NY2R_BOVIN STANDARD: PRT: 384 AA.
ID NY2R_BOVIN STANDARD: PRT: 384 AA.
AC P79113:
DE 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Amar D.A., Kolakowski L.F. Jr., Eadie D.M., Wong D.J., Ma Y.Y.,
RA Yang-Feng T.L., Thompson D.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHIKININS RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U50144; AAB40600.1;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 53 75 1 (POTENTIAL).
FT DOMAIN 76 85 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 86 107 2 (POTENTIAL).
FT DOMAIN 108 127 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 128 149 3 (POTENTIAL).
FT DOMAIN 150 169 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 170 190 4 (POTENTIAL).
FT DOMAIN 191 217 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 218 243 5 (POTENTIAL).
FT DOMAIN 244 271 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 272 294 6 (POTENTIAL).
FT DOMAIN 295 307 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 308 331 7 (POTENTIAL).
FT DOMAIN 332 384 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 126 206 BY SIMILARITY.
FT LIPID 345 345 PALMITATE (POTENTIAL).
SQ SEQUENCE 384 AA; 42943 MW; 468D19CBA8F29681 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 384;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRIFSKL 7
DB 243 TRIMSKL 249

RESULT 8
NY2R_MOUSE STANDARD: PRT: 385 AA.
ID NY2R_MOUSE STANDARD: PRT: 385 AA.

```

```

AC P97295;
DE 15-JUL-1998 (Rel. 36, Created)
DE 15-JUL-1998 (Rel. 36, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Brain;
RA MEDLINE=97071731; PubMed=8914576;
RA Nakamura M., Aoki Y., Hirano D.;
RT "Cloning and functional expression of a cDNA encoding a mouse type 2
RT neuropeptide Y receptor."
RL Biochim. Biophys. Acta 1284:134-137(1996).
CC -1- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHIKININS RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: D86238; BAA13050.1;
DR GCDDB: GCD:1146;
DR MGD: MG1:108418; NPY2r.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 53 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 54 76 1 (POTENTIAL).
FT DOMAIN 77 86 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 87 108 2 (POTENTIAL).
FT DOMAIN 109 128 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 129 150 3 (POTENTIAL).
FT DOMAIN 151 170 4 (POTENTIAL).
FT TRANSSEM 171 191 5 (POTENTIAL).
FT DOMAIN 192 218 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 219 244 6 (POTENTIAL).
FT DOMAIN 245 272 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 273 295 7 (POTENTIAL).
FT DOMAIN 296 308 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 309 332 7 (POTENTIAL).
FT DOMAIN 333 385 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 127 207 BY SIMILARITY.
FT LIPID 346 346 PALMITATE (POTENTIAL).
SQ SEQUENCE 385 AA; 43099 MW; 5AD7E4C9B7077085 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 385;
Best Local Similarity 85.7%; Pred. No. 51;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TRIFSKL 7
DB 244 TRIMSKL 250

RESULT 9
IAP1_DROME STANDARD: PRT: 385 AA.
ID IAP1_DROME STANDARD: PRT: 385 AA.

```

```

ID IAP1_DROME STANDARD: PRT: 438 AA.
AC Q24306;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apoptosis 1 inhibitor (inhibitor of apoptosis 1) (DIAP1) (Thread
DE protein).
GN IAP1 OR TH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE Imaginal disk;
RX MEDLINE=96128128; PubMed=8548811;
RA Hay B.A.; Wassarman D.A.; Rubin G.M.;
RT "Drosophila homologs of baculovirus inhibitor of apoptosis proteins
RT function to block cell death."
RL Cell 83:1253-1262(1995).
CC -1- FUNCTION: APOPTOTIC SUPPRESSOR. OVEREXPRESSION SUPPRESSES RPR AND
CC HID-DEPENDENT CELL DEATH IN THE EYE.
CC -1- SIMILARITY: BELONGS TO THE IAP FAMILY.
CC -1- SIMILARITY: CONTAINS 2 BIR REPEATS.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: LA9440; AAC1609.1; -.
DR HSSE: Q13490; 108H.
DR FLYBASE: FBgn0003691; th.
DR InterPro: IPR001370; BIR.
DR InterPro: IPR001841; znf_fing.
DR Pfam: PF00653; BIR: 2.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00238; BIR: 2.
DR SMART: SM00184; RING: 1.
DR PROSITE: PS01282; BIR_REPEAT_1; 2.
DR PROSITE: PS0143; BIR_REPEAT_2; 2.
DR PROSITE: PS00518; zf_RING_1; FALSE_NEG.
DR PROSITE: PS0089; zf_RING_2; 1.
DR Apoptosis; Zinc-finger; Repeat.
KW REPEAT 44 110 BIR 1.
FT REPEAT 226 293 BIR 2.
FT ZN_FING 391 426 RING-TYPE.
SQ SEQUENCE 438 AA; 48098 MW; A6C22C8EDF5A9F29 CRC64;

```

```

Query Match 84.8%; Score 28; DB 1; Length 438;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 TRIFSKL 7
DB 360 TRIFPKI 366

```

```

RESULT 10
XLIB_KLEPN STANDARD: PRT: 483 AA.
AC P29444;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Xylulose kinase (EC 2.7.1.17) (Xylulokinase).
GN XLIB.

```

```

OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1033-SP14 / KAY2026;
RX MEDLINE=92374988; PubMed=1324398;
RA Feldmann S.D.; Sahn H.; Sprenger G.A.;
RT "Cloning and expression of the genes for xylulose isomerase and
RT xylulokinase from Klebsiella pneumoniae 1033 in Escherichia coli
RT K12."
RL Mol. Gen. Genet. 234:201-210(1992).
CC -1- CATALYTIC ACTIVITY: ATP + D-xylulose -> ADP + D-xylulose 5-
CC phosphate.
CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCONOKINASE /
CC GLYCEROKINASE / XYLULOKINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X61059; CAA43390.1; -.
DR PIR: S25070; S25070.
DR InterPro: IPR000577; FGGY_kin.
DR Pfam: PF00370; FGGY; 1.
DR Pfam: PF02782; FGGY_C; 1.
DR PROSITE: PS00445; FGGY_KINASES_2; 1.
DR PROSITE: PS00933; FGGY_KINASES_1; 1.
KW Transferase; Kinase; Xylulose metabolism.
SQ SEQUENCE 483 AA; 51907 MW; 835FE494018FE872 CRC64;

```

```

Query Match 84.8%; Score 28; DB 1; Length 483;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 RIFSKL 7
DB 473 RIFPKI 478

```

```

RESULT 11
SCPL_HUMAN STANDARD: PRT: 976 AA.
AC Q15431; O14963;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein).
GN SCPL OR SCPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=97224467; PubMed=9119375;
RA Meuwissen R.L.J.; Meerts I.; Hoovers J.M.N.; Leschot N.J.;
RA Heyting C.;
RT "Human synaptonemal complex protein 1 (SCP1): isolation and
RT characterization of the cDNA and chromosomal localization of the
RT gene."
RL Genomics 39:377-384(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=98037449; PubMed=9371398;

```

RA Kondoh N., Nishina Y., Tsuchida J., Koga M., Tanaka H., Uchida K.,
 RA Inazawa J., Takeo M., Nozaki M., Nojima H., Matsumiya K., Namiki M.,
 RA Okuyama A., Nishimune Y.,
 RT Assignment of synaptonemal complex protein 1 (SCP1) to human
 RT chromosome 1p13 by fluorescence in situ hybridization and its
 RT expression in the testis.*;
 RL Cytogenet. Cell Genet. 78:103-104(1997).
 CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
 CC SYNAPTONEMAL COMPLEXES (SCS). FORMED BETWEEN HOMOLOGOUS
 CC CHROMOSOMES DURING MEIOTIC PROPHASE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
 CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.
 CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
 CC IS FOUND TOWARDS THE CENTRE OF THE SYNAPTONEMAL COMPLEX WHILE THE
 CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
 CC SYNAPTONEMAL COMPLEX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
 CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
 CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X95654; CAA64956.1; -;
 DR EMBL: D67035; BAA2586.1; -;
 DR MIM: 602162; -;
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 KM DNA-binding; Coiled coil.
 FT DOMAIN 12 100 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 107 798 COILED COIL (POTENTIAL).
 FT DOMAIN 117 120 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 679 682 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 880 883 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 961 969 ARG/LYS-RICH (BASIC).
 FT CONFLICT 46 46 L -> F (IN REF. 2).
 FT CONFLICT 106 106 F -> Y (IN REF. 2).
 FT CONFLICT 153 153 F -> C (IN REF. 2).
 FT CONFLICT 161 161 K -> T (IN REF. 2).
 FT CONFLICT 168 168 E -> D (IN REF. 2).
 FT CONFLICT 216 216 N -> S (IN REF. 2).
 FT CONFLICT 225 226 HG -> FE (IN REF. 2).
 FT CONFLICT 350 350 K -> N (IN REF. 2).
 FT CONFLICT 360 360 E -> D (IN REF. 2).
 FT CONFLICT 400 401 KN -> NY (IN REF. 2).
 FT CONFLICT 406 406 K -> I (IN REF. 2).
 FT CONFLICT 415 415 K -> T (IN REF. 2).
 FT CONFLICT 449 449 E -> D (IN REF. 2).
 FT CONFLICT 453 510 IQLTAITTSQYSEYKVDLTKTELENEK -> YSYCHYHKW
 FT CONFLICT 516 528 TVLPKRGQRPKLSKRE (IN REF. 2).
 FT CONFLICT 528 528 LITSHQNKLSLEK -> YFTLQQAAPPEN (IN REF. 2).
 FT CONFLICT 549 549 N -> I (IN REF. 2).
 FT CONFLICT 560 560 K -> T (IN REF. 2).
 FT CONFLICT 805 805 E -> D (IN REF. 2).
 FT CONFLICT 941 941 P -> S (IN REF. 2).
 SQ SEQUENCE 976 AA; 114069 MW; 8BA81D042AC2696B CRC64;

Query Match 84.8%; Score 28; DB 1; Length 976;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSRL 7
 DB 103 SRVFSKL 109

RESULT 12
 N145_YEAST
 ID N145_YEAST STANDARD; PRT; 1317 AA.
 AC P49687;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Nucleoporin NUP145 (Nuclear pore protein NUP145).
 GN NUP145 OR YG1092W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94320139; PubMed=8044840;
 RA Fabre E., Boelens W.C., Wimmer C., Mattaj J.W., Hurt E.C.;
 RT "Nup145p is required for nuclear export of mRNA and binds
 RT homopolymeric RNA in vitro via a novel conserved motif.";
 RL Cell 78:275-289(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 26109 / X2180;
 RX MEDLINE=94253246; PubMed=8195299;
 RA Wente S.R., Blobel G.;
 RT "NUP145 encodes a novel yeast glycine-leucine-phenylalanine-glycine
 RT (GLFG) nucleoporin required for nuclear envelope structure.";
 RL J. Cell Biol. 125:955-969(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97435481; PubMed=9290212;
 RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
 RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
 RT chromosome VII.";
 RL Yeast 13:1077-1090(1997).
 CC -1- FUNCTION: MAY PLAY A DIRECT ROLE IN NUCLEOCYTOPLASMIC EXPORT OF
 CC RNA AND ALSO IN PROTEIN IMPORT. BINDS HOMOPOLYMERIC RNA.
 CC -1- SUBCELLULAR LOCATION: Nuclear pore complex.
 CC -1- DOMAIN: CONTAINS G-L-F-G REPEATS.
 CC -1- SIMILARITY: BELONGS TO THE GLFG FAMILY OF NUCLEOPORINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X76557; CAA54057.1; -;
 DR EMBL: Z32672; CAA83584.1; -;
 DR EMBL: Z72614; CAA96798.1; -;
 DR SGD: S0003060; NUP145.
 KW Nuclear protein; Transport; Repeat; RNA-binding.
 FT CONFLICT 281 282 NA -> OR (IN REF. 2).
 FT CONFLICT 1142 1142 L -> S (IN REF. 2).
 FT CONFLICT 1310 1317 LMKCTYKI -> FEVYL (IN REF. 2).
 SQ SEQUENCE 1317 AA; 145660 MW; 59399D86B553030 CRC64;

Query Match 84.8%; Score 28; DB 1; Length 1317;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSRL 6
 DB 1068 TRVFSK 1073

RESULT 13
 CLIP_MOUSE
 ID CLIP_MOUSE STANDARD; PRT; 184 AA.

```

AC 09C36: Q9P7K7; Q9CTZ5;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE CAIL protein homolog.
GN CAIL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Stomach, and Tongue;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guslinch S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
RA Wyszew-Borla A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AK008990; BAB26010.1; -
DR EMBL: AK008622; BAB25784.1; -
DR EMBL: AK008641; BAB25801.1; -
DR EMBL: AK008647; BAB25805.1; -
DR EMBL: AK008722; BAB25856.1; -
DR EMBL: AK008745; BAB25872.1; -
DR EMBL: AK008933; BAB25975.1; -
DR EMBL: AK008956; BAB25988.1; -
DR EMBL: AK009145; BAB26103.1; -
DR EMBL: AK019050; BAB31525.1; -
DR CONFLICT 113 113
FT SEQUENCE 184 AA; 20134 MW; 288982F0404FEAB8 CRC64;
SQ

```

```

Query Match 81.8%; Score 27; DB 1; Length 184;
Best Local Similarity 83.3%; Pred. No. 43;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 TRFSK 6
DB 73 TRFSK 78

```

```

RESULT 14
Y860_RICPR STANDARD; PRT; 204 AA.
AC P11078;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Hypothetical protein RP860.
GN RP860.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OC NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RA Wood D.O., Marks G.L., Winkler H.H.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sierberitz-Ponten T., Alismar U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U02876; AAB81402.1; -
DR EMBL: AJ235273; CAA15284.1; -
DR InterPro: IPR002741; DUF56.
DR Pfam: PF01879; DUF56; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 204 AA; 23196 MW; E4CE96B478ECCCA6 CRC64;

```

```

Query Match 81.8%; Score 27; DB 1; Length 204;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 TRFSK 7
DB 63 TRFSK 69

```

```

RESULT 15
YC28_PORPU STANDARD; PRT; 237 AA.
AC P51342;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical transcriptional regulator Yc128.
GN YCF28.
OS Porphyra purpurea.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVONPORT;
RA Reith M.E., Munholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome.";
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- SIMILARITY: BELONGS TO THE CRP/FNR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U38604; AAC08228.1; -.
DR InterPro; IPR001808; HTH_CRP.
DR InterPro; IPR000595; CNMP_binding.
DR Pfam; PF00027; CNMP_binding; 1.
DR Pfam; PF00325; crp; 1.
DR SMART; SM00419; HTH_CRP; 1.
DR SMART; SM00100; CNMP; 1.
DR PROSITE; PS00042; HTH_CRP_FAMILY; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Chloroplast.
FT DNA BIND 188 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE ~237 AA; 27373 MW; CD8AAAF7C1DD4F333 CRC64;

Query Match 81.8%; Score 27; DB 1; Length 237;
Best Local Similarity 85.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRIFSKL 7
||| |||
Db 203 TRIISKL 209

Search completed: July 15, 2002, 13:23:31
Job time: 1447 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:22:39 ; Search time 172.49 Seconds
(without alignments)
7.020 Million cell updates/sec

Title: US-09-712-819A-2

Perfect score: 33
Sequence: 1 TRIPSKL 7

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	97.0	529	5 016232	016232 caenorhabd1
2	30	90.9	499	5 016230	016230 caenorhabd1
3	30	90.9	1106	16 09CBP3	09cbp3 mycobacteri
4	29	87.9	220	11 09CV09	09cv09 mus musculi
5	29	87.9	244	16 09A8K0	09a8k0 caulobacter
6	29	87.9	291	16 097KM7	097km7 clostridium
7	29	87.9	297	4 09BSH4	09bsh4 homo sapien
8	29	87.9	352	5 017134	017134 caenorhabd1
9	29	87.9	419	5 09BMX4	09bm4 cypocspor1
10	29	87.9	457	2 0930Z0	0930z0 chromatiu
11	29	87.9	579	10 09FN09	09fn09 arabidopsis
12	29	87.9	579	10 093V14	093v14 arabidopsis
13	29	87.9	590	10 098RV2	098rv2 guillardia
14	29	87.9	606	5 0961B6	0961b6 drosophila
15	29	87.9	1071	5 017405	017405 caenorhabd1
16	29	87.9	1522	10 0942K3	0942k3 oryza sativ

17	28	84.8	39	16 09JYS2	09jys2 neisseria m
18	28	84.8	76	11 070424	070424 rattus norv
19	28	84.8	149	2 09XBC4	09xbc4 amycolatops
20	28	84.8	179	16 09RZH3	09rzh3 delinococcus
21	28	84.8	181	2 09APN6	09apn6 pseudalter
22	28	84.8	187	16 051717	051717 borrelia bu
23	28	84.8	193	10 022830	022830 arabidopsis
24	28	84.8	295	16 098KJ9	098kj9 rhizobium 1
25	28	84.8	328	5 016953	016953 caenorhabd1
26	28	84.8	381	11 09ERC0	09erc0 rattus norv
27	28	84.8	422	5 020082	020082 caenorhabd1
28	28	84.8	438	5 09VUX5	09vux5 drosophila
29	28	84.8	545	12 09WAT6	09wat6 chuzan viru
30	28	84.8	580	10 0940F9	0940f9 arabidopsis
31	28	84.8	586	10 09SZF2	09szf2 arabidopsis
32	28	84.8	691	13 09PRH5	09prh5 oncorhynch
33	28	84.8	702	4 09HOK9	09hok9 homo sapien
34	28	84.8	744	10 048689	048689 arabidopsis
35	28	84.8	834	10 022187	022187 arabidopsis
36	28	84.8	1304	5 09BM81	09bm81 clonorchis
37	28	84.8	1304	5 09BM79	09bm79 clonorchis
38	28	84.8	1311	5 09BM80	09bm80 clonorchis
39	28	84.8	1491	5 09NBD7	09nbd7 drosophila
40	28	84.8	1492	5 09NL57	09nl57 drosophila
41	27	81.8	156	17 029614	029614 archaeoglob
42	27	81.8	161	11 09D9A9	09d9a9 mus musculi
43	27	81.8	252	12 09DYY7	09dyy7 pluteella xy
44	27	81.8	258	4 09H737	09h737 homo sapien
45	27	81.8	338	4 0960M7	0960m7 homo sapien

ALIGNMENTS

RESULT 1	
ID 016232	PRELIMINARY; PRT; 529 AA.
AC 016232;	
DT 01-JAN-1998 (TREMBLrel. 05, Created)	
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)	
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE HYPOTHETICAL 61.2 KDA PROTEIN.	
GN F07G11.3.	
OS Caenorhabditis elegans.	
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OC Rhabditidae; Pelodierinae; Caenorhabditis.	
OX NCBI_TaxID=6239;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-BRISTOL N2;	
RX MEDLINE-99069613; PubMed=9851916;	
RA None;	
RT "Genome sequence of the nematode C. elegans: a platform for	
RT investigating biology. The C. elegans Sequencing Consortium.;"	
RL Science 282:2012-2018(1998).	
RN [2]	
RP SEQUENCE FROM N.A.	
RC STRAIN-BRISTOL N2;	
RA Sammons L., Wohlmann P., Sansone J.;	
RL "The sequence of C. elegans cosmid F07G11.1;"	
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.	
RN [3]	
RP SEQUENCE FROM N.A.	
RC STRAIN-BRISTOL N2;	
RA Waterston R.;	
RL "Direct Submission.;"	
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AF016419; AAG24048.1;	
DR InterPro; IPR002875; DUF23.	
DR Pfam; PF01697; DUF23; 1.	
DR Prodom; PD004153; DUF23; 1.	
KW Hypothetical protein.	
SQ SEQUENCE 529 AA; 61221 MW; CEAAF583B95558B0 CRC64;	

Query Match 97.0%; Score 32; DB 5; Length 529;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
||:|||||
DB 440 TRVFSKL 446

RESULT 2
016230 PRELIMINARY; PRT; 499 AA.

AC 016230;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOPHYSICAL 58.2 KDA PROTEIN.
GN F07G11.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=9069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sammons L., Mohldmann P., Sansone J.;
RT "The sequence of C. elegans cosmid F07G11.";
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submision.";
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF016418; AAC24051.1;
DR InterPro: IPR002875; DUF23.
DR Pfam: PF01697; DUF23; 1.
DR ProDom: PD004153; DUF23; 1.
KW Hypothetical protein.
SO SEQUENCE 499 AA; 58227 MW; 53F8F7E6871FB9F5 CRC64;

Query Match 90.9%; Score 30; DB 5; Length 499;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
||:|||||
DB 410 TRVFSKL 416

RESULT 3
09CBP3 PRELIMINARY; PRT; 1106 AA.

AC 09CBP3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE POSSIBLE TRANSCRIPTIONAL REGULATOR.
GN ML1753.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=2118732; PubMed=11234002;
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: BELONGS TO THE LUXR/UHPA FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC EMBL: AL583923; CAC30706.1; -.
DR Leproma; ML1753; -.
DR InterPro: IPR000767; Disease_resist.
DR InterPro: IPR000792; HTL_LuxR.
DR Pfam: PF00196; GERE; 1.
DR PRINTS: PR00364; DISEASERSIST.
DR PRINTS: PR00038; HTMLUXR.
DR SMART: SM00421; HTH_LuxR; 1.
KW Complete proteome; DNA-binding; Transcription regulation.
SO SEQUENCE 1106 AA; 119374 MW; 75368FB4C7C99DE2 CRC64;

Query Match 90.9%; Score 30; DB 16; Length 1106;
Best Local Similarity 85.7%; Pred. No. 1,7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
||:|||||
DB 1084 TRIVSKL 1090

RESULT 4
09CV09 PRELIMINARY; PRT; 220 AA.

AC 09CV09;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 2310066118RIK PROTEIN (FRAGMENT).
GN 2310066118RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli K., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carinacci P., de Bonaldi M.F.,
RA Brownstein M.J., Bolt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welz C., Whitaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawai J.H., Kohsaki S.,

RA Hayashizaki Y.:
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK010072; BAB26682.1; -
 DR MGI: 1917457; 2310066118Rik.
 DR InterPro: IPR002876; DUF28.
 DR Prodom: PD004323; DUF28.1.
 DR NON_TER 220
 FT SEQUENCE 220 AA; 23862 MW; F78AC7088AAB4BB6 CRC64;

Query Match
 Best Local Similarity 87.9%; Score 29; DB 11; Length 220;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSKL 7
 :|||||
 DB 75 SRIFSKL 81

RESULT 5
 Q9A8K0 PRELIMINARY; PRT; 244 AA.
 AC Q9A8K0;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN CCI1353.
 GN CCI1353.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohra N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phade N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Dworkin A.S., Gwin M.L., Haft D.H.,
 RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005810; AAK23334.1; -
 DR TIGR: CCI1353; -
 DR InterPro: IPR001601; Meth-transf.
 DR InterPro: IPR000051; SAM_bind.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 244 AA; 26206 MW; DE0382178696DEFL CRC64;

Query Match
 Best Local Similarity 87.9%; Score 29; DB 16; Length 244;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSK 6
 :|||||
 DB 69 TRIFSK 74

RESULT 6
 Q97KW7 PRELIMINARY; PRT; 291 AA.
 AC Q97KW7;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PHOSPHATIDYLKERINE DECARBOXYLASE.
 GN CAC0799.
 OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RX MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hilti J., Wolf Y.I.,
 RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Souaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007595; AAK78775.1; -
 DR InterPro: IPR003817; PS_Dcarboxylase.
 DR Pfam: PF02666; PS_Dcarboxylase; 1.
 KW Complete proteome.
 SQ SEQUENCE 291 AA; 34070 MW; 6CC2F1BCC87489D CRC64;

Query Match
 Best Local Similarity 87.9%; Score 29; DB 16; Length 291;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSKL 7
 :|||||
 DB 43 TRIFSKI 49

RESULT 7
 Q9BSH4 PRELIMINARY; PRT; 297 AA.
 AC Q9BSH4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE SIMILAR TO RIKEN CDNA 2310066118 GENE (UNKNOWN) (PROTEIN FOR
 MGC:13286).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA; CHORIOCARCINOMA;
 RA Strusberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN; NEUROBLASTOMA;
 RA Strusberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC005049; AAH05049.1; -
 DR EMBL: BC007744; AAH07744.1; -
 DR InterPro: IPR002876; DUF28.
 DR Pfam: PF01709; DUF28.1.
 DR Prodom: PD004323; DUF28.1.
 SQ SEQUENCE 297 AA; 32477 MW; 079C4716F32EE6FF CRC64;

Query Match
 Best Local Similarity 87.9%; Score 29; DB 4; Length 297;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSKL 7
 :|||||
 DB 78 SRIFSKL 84

RESULT 8
 O17134 PRELIMINARY; PRT; 352 AA.
 ID O17134

AC 017134;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE HYPOTHEICAL 40.6 KDA PROTEIN.
GN F31F4.14.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RT "The sequence of C. elegans cosmid F31F4.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RT "Direct Submission.";
RL EMBL: AF024503; AAC24092.1;
DR EMBL: AF024503; AAC24092.1;
DR InterPro: IPR003839; DUF215.
DR Pfam: PF02688; DUF215; 1.
KW Hypothetical protein.
SQ SEQUENCE 352 AA; 40599 MW; BADAA1293C9B7585 CRC64;

Query Match 87.9%; Score 29; DB 5; Length 352;
Best Local Similarity 85.7%; Pred. No. 94;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
DB 48 SRIFSKL 54

RESULT 9
O9BMX4 PRELIMINARY; PRT; 419 AA.
AC O9BMX4:
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE URIDINE KINASE/URACIL PHOSPHORIBOSYLTRANSFERASE.
GN UKRPR1.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RA Woods K.M., Upton S.J.;
RT "Sequence of the uridine kinase/uracil phosphoribosyltransferase Gene
from Cryptosporidium parvum.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF298155; AAC53652.1;
DR HSSP: Q26998; IBD3.
DR InterPro: IPR001324; PRK.
DR InterPro: IPR000764; Uridine_kin.
DR Pfam: PF00485; PRK; 1.
DR PRINTS: PR00988; URIDINKINASE.
KW Transferase; Kinase; Glycosyltransferase.
SQ SEQUENCE 419 AA; 47640 MW; 321BC8B143B545CF CRC64;

Query Match 87.9%; Score 29; DB 5; Length 419;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
DB 305 TRVFTKL 311

RESULT 11
O9FNO9 PRELIMINARY; PRT; 579 AA.
AC O9FNO9:
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PHOTORECEPTOR-INTERACTING PROTEIN-LIKE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
Sequence features of the regions of 1,191,918 bp covered by seventeen

RT physically assigned P1 clones."
RL DNA Res. 4:401-414(1997).
DR EMBL; AB007645; BAB09029.1; -
DR InterPro; IPR000210; BTB_P0Z.
DR InterPro; IPR004249; NPH3.
DR Pfam; PR03000; NPH3; 1.
DR SMART; SM00225; BTB; 1.
DR PROSITE; P550097; BTB; 1.
KW Receptor.
SQ SEQUENCE 579 AA; 64765 MW; FD24DB00C95497B1 CRC64;

Query Match
Best Local Similarity 87.9%; Score 29; DB 10; Length 579;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
Db 531 SRIFSKL 537

RESULT 12
O93Y14 PRELIMINARY; PRT; 579 AA.
AC O93Y14;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE PHOTORECEPTOR-INTERACTING PROTEIN-LIKE.
GN AT5G67440 OR K8K14.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamaya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shino P., Yamada K., Yoshizaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054694; AAK96885.1; -
KW Receptor.
SQ SEQUENCE 579 AA; 64764 MW; F8EDFA8DAADC3E93 CRC64;

Query Match
Best Local Similarity 87.9%; Score 29; DB 10; Length 579;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
Db 531 SRIFSKL 537

RESULT 13
O98RV2 PRELIMINARY; PRT; 590 AA.
AC O98RV2;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE HYPOTHETICAL 71.5 KDA PROTEIN ORF590 FROM CHROMOSOME 1.
GN ORF590.
OS Gulliardia theta (Cryptomonas phi).
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Gulliardia.
OX NCBI_TaxID=55529;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21223349; Pubmed=11323671;

RA Douglas S., Zauner S., Fraunholz M., Beaton M., Penny S., Deng L.T.,
Wu X., Reith M., Cavallier-Smith T., Maier U.G.;
RT "The highly reduced genome of an enslaved algal nucleus."
RL Nature 410:1091-1096(2001).
DR EMBL; AF165818; AAK39848.1; -
KW Hypothetical protein.
SQ SEQUENCE 590 AA; 71520 MW; 4FDC746B1A7330FD CRC64;

Query Match
Best Local Similarity 87.9%; Score 29; DB 10; Length 590;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
Db 203 SRIFSKL 209

RESULT 14
O961B6 PRELIMINARY; PRT; 606 AA.
ID O961B6
AC O961B6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE LD24529P.
GN TAR60.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=7, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Paclele J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051702; AAK93126.1; -
SQ SEQUENCE 606 AA; 65654 MW; 9D93C3496A23C0A3 CRC64;

Query Match
Best Local Similarity 87.9%; Score 29; DB 5; Length 606;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6
Db 351 TRIFSK 356

RESULT 15
O17405 PRELIMINARY; PRT; 1071 AA.
ID O17405
AC O17405;
DT 01-JAN-1999 (TREMBLrel. 09, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 120.6 KDA PROTEIN AC3.5 IN CHROMOSOME V.
GN AC3.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Mcmurray A.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: WEAK, TO PEPTIDASE FAMILY M1.

DR EMBL: Z71177; CAA94872.1; -.
 DR MEROPS: M01.UNW; -.
 DR WormPep; AC3.5; CE05135.
 DR InterPro; IPR001930; Aladiptase.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF01433; Peptidase_M1; 1.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 81 101 POTENTIAL.
 SQ SEQUENCE 1071 AA: 120605 MW: 09733f0A0ECF231F CRC64;

Query Match 87.9%; Score 29; DB 5; Length 1071;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TRIFSKL 7
 |||||
 Db 788 TRIFKKL 794

Search completed: July 15, 2002, 13:22:40
 Job time: 1476 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:57:54 ; Search time 228.39 Seconds
(without alignments)
3.404 Million cell updates/sec

Title: US-09-712-819A-2
Perfect score: 33
Sequence: 1 TRIPSKL 7

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	87.9	297	22	AAU27642	Human protein AFP5
2	29	87.9	297	22	AAU39534	Human polypeptide
3	29	87.9	303	22	AAU41320	Human polypeptide
4	29	87.9	591	15	AAU56489	TATA-binding prote
5	29	87.9	591	17	AAU06080	Drosophila TATA-bi
6	29	87.9	591	18	AAU25026	TATA-binding prote
7	29	87.9	592	22	ABU71746	Drosophila melanog
8	28	84.8	128	22	AAU06882	Human polypeptide
9	28	84.8	170	21	AAU61781	Arabidopsis thalia
10	28	84.8	166	20	AAU19828	B. burgdorferi ant
11	28	84.8	207	22	AAU90104	C. glutamicum prote

12	28	84.8	381	16	AAU78271	Human hippocampal
13	28	84.8	381	16	AAU78272	Rat hippocampal ne
14	28	84.8	381	16	AAU78273	Rat hippocampal ne
15	28	84.8	381	19	AAU41710	Rhesus monkey neur
16	28	84.8	381	21	AAU56888	Human Y2 receptor
17	28	84.8	381	22	AAU07921	Human neurotrophin
18	28	84.8	381	22	AAU07957	Human neurotrophin
19	28	84.8	381	22	AAU06689	Human neurotrophin
20	28	84.8	381	22	AAU85109	Human neurotrophin
21	28	84.8	381	22	AAU85120	Human neurotrophin
22	28	84.8	438	22	AAU85185	Drosophila melanog
23	28	84.8	438	22	AAU85186	Drosophila melanog
24	28	84.8	438	22	AAU85187	Drosophila wild-ty
25	28	84.8	438	22	AAU85188	Drosophila wild-ty
26	28	84.8	438	22	AAU85189	Drosophila mutant
27	28	84.8	438	22	AAU85190	Drosophila mutant
28	28	84.8	438	22	AAU85191	Drosophila mutant
29	28	84.8	438	22	AAU85192	Drosophila mutant
30	28	84.8	438	22	AAU85193	Drosophila mutant
31	28	84.8	438	22	AAU85194	Drosophila mutant
32	28	84.8	438	22	AAU85195	Drosophila mutant
33	28	84.8	604	22	AAU51655	Protonibacterium
34	28	84.8	976	22	AAU66581	Human SCP-1 mutain
35	28	84.8	1492	22	AAU31597	Amino acid sequenc
36	27	81.8	49	21	AAU39989	Arabidopsis thalia
37	27	81.8	50	22	AAU21482	Human novel foetal
38	27	81.8	65	22	AAU02168	Human polypeptide
39	27	81.8	67	22	AAU03192	Human polypeptide
40	27	81.8	136	21	AAU24639	Plant SPF encoded
41	27	81.8	137	21	AAU44092	Human cancer assoc
42	27	81.8	138	21	AAU24638	Plant SPF encoded
43	27	81.8	140	21	AAU24637	Human cancer assoc
44	27	81.8	162	21	AAU43503	Arabidopsis thalia
45	27	81.8	181	21	AAU14002	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AAU27642	
ID	AAU27642 standard; Protein: 297 AA.
AC	AAU27642;
XX	
DT	18-DEC-2001 (first entry)
DE	Human protein AFP50993.
XX	
KW	Human; AFR; cytosolic; Marfan's syndrome; thrombocytopenia;
KW	leukaemia; porphyria; Gilles De La Tourette's syndrome; immunogen;
KW	squamous cell carcinoma; diabetes mellitus; Grave's disease;
KW	colon cancer; Alzheimer's disease; epiphyseal dysplasia.
XX	
OS	Homo sapiens.
XX	
PN	WO200166748-A2.
XX	
PD	13-SEP-2001.
XX	
PF	05-MAR-2001; 2001WO-US07192.
XX	
PR	03-MAR-2000; 2000US-187221P.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PI	Conklin DC, Presnell SR, Adler DA;
XX	
DR	WPI; 2001-589943/66.
XX	
DR	N-PSDB; AAS44543.
XX	
PT	Novel AFP polypeptides and polynucleotides, useful for diagnostic and therapeutic purposes, in cancer therapy and for screening modulator

PT compounds -

XX Claim 2: Page 135-136; 220pp; English.

XX

CC The invention relates to novel human AFP proteins (not defined) and the nucleic acids that encode them. AFP proteins are useful as standards in assays of protein and amino acid supplements, including hydrolyses. CC settings, as protein and amino acid supplements, including hydrolyses. CC The nucleic acids are useful for radiation hybrid mapping. Secretary CC fusion proteins of AFP are useful in cancer therapy, for enhancing in vitro cytotoxicity, for enhancing in vivo killing of target tissues, and for targeted cells or tissue inhibition or ablation. Anti-AFP antibodies are useful for isolating target polypeptides by affinity purification, in CC diagnostic assays for determining circulating or localised levels of target polypeptides, for tissue typing, for cell sorting, for screening CC expression libraries, for generating anti-idiotypic antibodies, and as CC neutralising antibodies or as antagonists to block protein activity in CC vitro and in vivo. AFP proteins and nucleic acids may be used to CC diagnose or treat (e.g. by gene therapy) diseases associated with the CC malfunction of the AFP e.g. AFP166924 and Marfan's syndrome, CC AFP576853/AFP39158 and thrombocytopaenia, leukaemia, porphyria, Gilles De La Tourette's syndrome, AFP652829 and squamous cell carcinoma, diabetes CC mellitus, Grave's disease, AFP664311 and colon cancer. AFP251034 and CC Alzheimer's disease and AFP686580 and epiphyseal dysplasia. Many more CC examples of disease are given in the specification. The present CC sequence represents an AFP of the invention.

XX

SQ Sequence 297 AA;

Query Match 87.9%; Score 29; DB 22; Length 297;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
:|||||
Db 78 srtfskl 84

RESULT 2
AAM39534
ID AAM39534 standard; Protein: 297 AA.
XX
AC AAM39534;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2679.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW Leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0661191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX

PA (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Weinman F, Xu C, Xue AD, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
DR N-PSDB: AA158690.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
PS Example 4; SEQ ID NO 2679; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, CC immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide CC of the invention may be used to treat diseases of the peripheral nervous CC system, such as peripheral nervous injuries, peripheral neuropathy and CC localised neuropathies and central nervous system diseases, such as CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the CC utilisation of the activities such as: Immune system suppression, CC Activin/Inhibin activity, chemotactic/chemokine activity, haemostatic CC and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukaemias and CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed CC specification.

XX

SQ Sequence 297 AA;

Query Match 87.9%; Score 29; DB 22; Length 297;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
:|||||
Db 78 srtfskl 84

RESULT 3
AAM41320
ID AAM41320 standard; Protein: 303 AA.
XX
AC AAM41320;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6251.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW Leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
XX

```

PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HXSE-) HYSEQ INC.
PA
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AA160476.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 6251; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AA038642-AA042213) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention. A composition containing a polypeptide or polynucleotide
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
XX Sequence 303 AA;
SO

```

PI	Comai L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;
PI	Tjian R, Wang E, Weinzierl ROJ;
XX	
DR	WPI: 1994-264019/32.
XX	
DR	N-PSDB; AAO70727.
XX	
PT	TATA-binding protein associated protein factors - and
PT	corresponding nucleotide sequence and deriv. antibodies, useful
PT	in screening, diagnostics and therapeutics
XX	
PS	Disclosure: Page 85-88; 180pp; English.
XX	
CC	The TATA-binding protein associated factor dTAFII60 (including
CC	specific antibodies and fusion products) are used in drug screening,
CC	diagnostics and therapeutics. They are used in the development of
CC	specific biochemical assays for screening compounds that agonise or
CC	antagonise selected transcription factors involved in regulating
CC	gene expression associated with human pathology.
XX	
SQ	Sequence 591 AA;
QY	1 TRIPSK 6
Db	337 trlfsk 342
RESULT 5	
AAW06080	
ID	AAW06080 standard; Protein; 591 AA.
XX	
AC	AAW06080;
DT	27-JAN-1997 (first entry)
XX	
DE	Drosophila TATA-binding protein associated factor dTAFII60 protein.
XX	
KW	Drosophila: TATA-binding protein; TBP associated factor; TFIID;
KW	RNA polymerase II; transcription; messenger RNA; nuclear fraction;
KW	holoenzyme; lambda-gt11; expression library.
XX	
OS	Drosophila melanogaster.
XX	
PN	US534410-A.
XX	
PD	09-JUL-1996.
XX	
PF	28-JAN-1993; 93US-0013412.
XX	
PR	28-JAN-1994; 94US-0188582.
PR	28-JAN-1993; 93US-0013412.
PR	30-JUN-1993; 93US-0087119.
XX	
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Comai L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;
PI	Tjian R, Wang E, Weinzierl ROJ;
XX	
XX	WPI: 1996-333245/33.
DR	N-PSDB; AAT42213.
XX	
PT	Screen for cpds. that bind human TATA-binding protein associated
PT	factor - by testing ability to bind to polypeptide fragments of the
PT	factor, useful as (ant)agonists of transcription factors involved in
PT	disease.
XX	
PS	Examples: Column 53-58; 86pp; English.
XX	
CC	This is the amino acid sequence encoding the Drosophila TATA-binding

CC protein (TBP) associated factor (TAF) designated TAFII60. The protein
 CC is a component of the TFIID fraction required for reconstituting RNA
 CC polymerase II in vitro transcription activity. The encoded protein
 CC has an estimated mol. wt. of 60 kd by SDS-PAGE.
 CC The invention relates to purified proteins involved in transcription
 CC by RNA polymerase II, the RNA polymerase which transcribes messenger
 CC RNA. RNA polymerase II transcription proceeds in vitro upon addition
 CC of several nuclear fractions designated TFIIA, B, D, E, F, H, I and J
 CC to RNA polymerase II holoenzyme. Fraction TFIID has been shown to
 CC contain a TBP and other TAFs. Purification of TFIID and separation of
 CC its components reveals 7 proteins ranging in size from 30-250 kd.
 CC Serum raised against the TFIID fraction allowed cloning of the corresp.
 CC genes from lambda-gt11 expression libraries.

CC Sequence 591 AA;

Query Match 87.9%; Score 29; DB 17; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6

Db 337 trlfsk 342

RESULT 6

AAW25026 AAW25026 standard; Protein; 591 AA.

AC AAW25026;

DT 08-OCT-1997 (first entry)

DE TATA-binding protein associated factor, dTAFII60.

KM TATA-binding protein associated factor; TAF; nuclear protein;

KW RNA polymerase transcription; TATA-binding protein; TBP;

OS Drosophila sp.

PN US5537686-A.

PD 10-JUN-1997.

PE 28-JAN-1993; 93US-0013412.

PR 28-JAN-1994; 94US-0188582.

PR 28-JAN-1993; 93US-0013412.

PR 30-JUN-1993; 93US-0087119.

PR 09-MAY-1996; 96US-0646715.

PA (REGC) UNITV CALIFORNIA.

PI Comai L, Dynlacht BD, Hoey T, Ruppert S, Tanese N;

PI Tjian R, Wang E, Weinzierl ROJ;

DR WPI; 1997-319113/29.

DR N-PSDB; AAT79602.

XX Nucleic acids encoding human TATA-binding protein associated factor

PT (TAF) peptide(s) - for production of recombinant peptide(s), used

PT for modulating transcription of TAFs

CC Example 1; Column 57-60; 86pp; English.

CC AAW25026 represents TATA-binding protein associated factor (TAF)

CC polypeptide, dTAFII60 (mol. weight 60kd). TAF peptides derived

CC from dTAFII30 alpha, dTAFII30 beta, dTAFII40, dTAFII60, dTAFII80,

CC dTAFII10, dTAFII150, and dTAFII250, their human equivalents and

CC nucleic acids encoding them, are used to modulate transcription,

CC including transcription initiation. TAFs are nuclear proteins involved

CC in RNA polymerase I, II and III transcription. The peptides act by
 CC binding to a different TAF, an activator, or TBP (TATA-binding protein)
 CC or competitively inhibiting association of a TAF domain with another
 CC compound, typically a protein like TBP or another TAF, an activator,
 CC or DNA.

CC Sequence 591 AA;

Query Match 87.9%; Score 29; DB 18; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6

Db 337 trlfsk 342

RESULT 7

ABB71746 ABB71746 standard; Protein; 592 AA.

AC ABB71746;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 42030.

KM Drosophila: developmental biology; cell signalling; insecticide;

KW Pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PR 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-6556860/75.

DR N-PSDB; ABL15849.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

XX Disclosure; SEQ ID NO 42030; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins

CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct-sequences.

XX Sequence 592 AA;

Query Match 87.9%; Score 29; DB 22; Length 592;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6
| | | | |
Db 337 trlfsk 342

RESULT 8

AAO06982
ID AAO06982 standard; Protein: 128 AA.

XX AAO06982;

DT 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 20874.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;
KM nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI86913.

PT Isolated nucleic acids and polypeptides, useful for preventing
disorders - diagnosing and treating e.g. leukaemia, inflammation and immune

PS Claim 20; SEQ ID NO 20874; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 128 AA;

Query Match 84.8%; Score 28; DB 22; Length 128;

Best Local Similarity 83.3%; Pred. No. 85;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSK 6
| | | | |

Db 15 trlfsk 20

RESULT 9

AAG61781

ID AAG61781 standard; Protein: 170 AA.

XX AAG61781;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 80243.

XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127452.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132487.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134221.

XX 14-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 16-JUN-1999; 99US-0139453.

XX 17-JUN-1999; 99US-0139454.

XX 18-JUN-1999; 99US-0139455.

XX 18-JUN-1999; 99US-0139456.

XX 18-JUN-1999; 99US-0139457.

XX 18-JUN-1999; 99US-0139458.

XX 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150565.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156459.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 84.8%; Score 28; DB 21; Length 170;
Best Local Similarity 71.4%; Pred. No. 1,1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TRIRSKL 7
Db 128 srfsk1 134
RESULT 10
AAV19828
AAV19828 standard; Protein; 186 AA.
XX
AC AAV19828;
XX
DT 19-JUL-1999 (first entry)
XX

DE B. burgdorferi antigenic protein, f82.aa.
XX Antigenic protein; vaccine; Lyme disease; infection; detection.
KM
XX
OS Borrelia burgdorferi.
XX
PN WO9859071-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US127718.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX
DR WPI: 1999-189980/16.
XX
DR N-PSDB: AAX61525.
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the diagnosis, prevention and treatment of diseases
PT caused by Borrelia, particularly Lyme disease
XX
PS Claim 12; Page 82; 275bp; English.
XX
CC This sequence represents a Borrelia burgdorferi (Bb) protein of the
CC invention, which is suitable for use in a vaccine. The Bb polypeptides
CC can be used in vaccines for eliciting protective antibodies to members of
CC the Borrelia genus, particularly for the use against Lyme disease in
CC humans and animals. They can be used for preventing or attenuating an
CC infection caused by a member of the Borrelia genus. The products can also
CC be used for detection of members of the Borrelia genus.
XX
SQ Sequence 186 AA;

Query Match 84.8%; Score 28; DB 20; Length 186;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRFSK 6
DB 2 trvfsk 7

RESULT 11
ID AAG90104 standard; Protein; 207 AA.
XX
AC AAG90104;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamylc protein fragment SEQ ID NO: 3858.
XX
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PR 18-DEC-2000; 2000EP-0127688.
XX
PR 16-DEC-1999; 99JP-0377484.

PR 07-APR-2000; 2000JP-0159162.
PR 03-AUG-2000; 2000JP-0280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
DR WPI: 2001-376931/40.
DR N-PSDB: AAH65323.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analysing
PT expression profile or pattern of a gene and identifying homologous gene
PT
XX
PS Claim 17; SEQ ID NO: 3858; 246pp + Sequence listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and
CC analysing the expression profile or expression pattern of a gene derived
CC from Coryneform bacterium, and identifying a homologue of a gene derived
CC from Coryneform bacterium. Coryneform bacteria are useful for producing
CC amino acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described
CC in the exemplification of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 207 AA;

Query Match 84.8%; Score 28; DB 22; Length 207;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRFSK 6
DB 127 trvfsk 132

RESULT 12
ID AAR78271 standard; Protein; 381 AA.
XX
AC AAR78271;
XX
DT 06-MAR-1996 (first entry)
XX
DE Human hippocampal neuropeptide Y2 receptor (clone CG-13).
XX
XX Hippocampus; brain; neuropeptide; peptide; hypertension; pain;
KW gastrointestinal disorder; diagnosis; sleeping disorder; epilepsy;
KW memory loss; diarrhoea; nasal congestion.
XX
OS Homo sapiens.
XX
PN WO9521245-A1.
XX
PD 10-AUG-1995.
XX
PR 03-FEB-1995; 95WO-US01469.
XX
PR 03-FEB-1994; 94US-0192288.
XX
PA (SYNA-) SYNAPTIC PHARM CORP.
XX
PI Branchek T, Gerald C, Walker MW, Weinschank R;
XX
DR WPI: 1995-283765/37.

DR N-PSDB; AAQ95034.
 XX
 PT Human and rat Y2 receptor DNA and protein - useful in diagnosis and
 PT treatment of e.g. cognitive and gastrointestinal disorder(s),
 PT hypertension and pain
 XX
 PS Disclosure: Page 115-117; 193pp; English.
 XX
 CC This protein, which may be expressed recombinantly in NIH3T3 cells,
 CC designated N-hy2-5 (ATCC CRL-11825) cells, using plasmid pCEXV-hy2
 CC (ATCC 75659), may be useful in the treatment of several
 CC pathophysiological conditions including cognitive, gastrointestinal
 CC or sleeping disorders, epilepsy, hypertension, memory loss, diarrhoea,
 CC nasal congestion or pain.
 XX
 SQ Sequence 381 AA;

Query Match 84.8%; Score 28; DB 16; Length 381;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 |||:|
 DB 240 trlwskl 246

RESULT 13
 AAR78272
 ID AAR78272 standard; Protein: 381 AA.

AC AAR78272;

DT 06-MAR-1996 (first entry)

DE Rat hippocampal neuropeptide Y2 receptor (clone Rs25a).

XX Hippocampus; brain; neuropeptide; peptide; hypertension; pain;

KW gastrointestinal disorder; diagnosis; sleeping disorder; epilepsy;

KW memory loss; diarrhoea; nasal congestion.

OS Rattus rattus.

PN W09521245-A1.

PD 10-AUG-1995.

PF 03-FEB-1995; 95MO-US01469.

PR 03-FEB-1994; 94US-0192288.

PA (SYNA-) SYNAPTIC PHARM CORP.

PI Branchek T, Gerald C, Walker MW, Weinschank R;

DR WPI: 1995-283765/37.

DR N-PSDB; AAQ95035.

XX Human and rat Y2 receptor DNA and protein - useful in diagnosis and

PT treatment of e.g. cognitive and gastrointestinal disorder(s),

PT hypertension and pain

XX Disclosure: Page 119-120; 193pp; English.

CC This protein, which may be expressed recombinantly in NIH3T3 cells,
 CC designated N-hy2-5 (ATCC CRL-11825) cells, using plasmid pCEXV-r2a
 CC (ATCC 97035), may be useful in the treatment of several
 CC pathophysiological conditions including cognitive, gastrointestinal
 CC or sleeping disorders, epilepsy, hypertension, memory loss, diarrhoea,
 CC nasal congestion or pain.
 XX

SQ Sequence 381 AA;

Query Match 84.8%; Score 28; DB 16; Length 381;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 |||:|
 DB 240 trlwskl 246

RESULT 14
 AAR78273
 ID AAR78273 standard; Protein: 381 AA.

AC AAR78273;

DT 06-MAR-1996 (first entry)

DE Rat hippocampal neuropeptide Y2 receptor (clone Rs26a).

XX Hippocampus; brain; neuropeptide; peptide; hypertension; pain;

KW gastrointestinal disorder; diagnosis; sleeping disorder; epilepsy;

KW memory loss; diarrhoea; nasal congestion.

OS Rattus rattus.

PN W09521245-A1.

PD 10-AUG-1995.

PF 03-FEB-1995; 95MO-US01469.

PR 03-FEB-1994; 94US-0192288.

PA (SYNA-) SYNAPTIC PHARM CORP.

PI Branchek T, Gerald C, Walker MW, Weinschank R;

DR WPI: 1995-283765/37.

DR N-PSDB; AAQ95036.

XX Human and rat Y2 receptor DNA and protein - useful in diagnosis and

PT treatment of e.g. cognitive and gastrointestinal disorder(s),

PT hypertension and pain

XX Disclosure: Page 122-124; 193pp; English.

CC This protein, which may be expressed recombinantly in NIH3T3 cells,
 CC designated N-hy2-5 (ATCC CRL-11825) cells, using plasmid pCEXV-r2b
 CC (ATCC 97036), may be useful in the treatment of several
 CC pathophysiological conditions including cognitive, gastrointestinal
 CC or sleeping disorders, epilepsy, hypertension, memory loss, diarrhoea,
 CC nasal congestion or pain.
 XX

SQ Sequence 381 AA;

Query Match 84.8%; Score 28; DB 16; Length 381;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
 |||:|
 DB 240 trlwskl 246

RESULT 15
 AAM41710
 ID AAM41710 standard; Protein: 381 AA.

AC AAM41710;

DT 22-JUN-1998 (first entry)

XX Rhesus monkey neuropeptide Y2 receptor.
 DE
 XX
 KM Neuropeptide Y2 receptor; rhesus monkey; neurotransmitter; angina;
 KM pain; renal failure; neurodegeneration; metabolic disorder;
 KM obesity; inflammation; asthma; vasospasm; heart failure; shock;
 KM cardiac hypertrophy; hypertension; myocardial infarction;
 KM arrhythmia; epilepsy; cerebral infarction; stroke; haemorrhage;
 KM cerebral vasospasm; depression; anxiety; schizophrenia; dementia;
 KM incontinence; Crohn's disease; anorexia; bronchoconstriction;
 KM therapy.
 XX
 OS Macaca mulatta.
 XX
 PN MO9748406-A1.
 XX
 PD 24-DEC-1997.
 XX
 PF 17-JUN-1997; 97WO-US10048.
 XX
 PR 17-JUN-1996; 96US-0019809.
 XX
 PA (E.L.L.) LILLY & CO ELI.
 XX
 PI Baez M, Yang P;
 XX
 DR WPI: 1998-062842/06.
 XX N-PSDB: AAV05530-31.
 PT Rhesus monkey neuropeptide Y2 receptor - useful to develop
 PT treatments for, e.g. angina, renal failure, neurodegeneration, pain,
 PT obesity, metabolic disorder, inflammation or asthma
 XX
 PS Claim 1: Page 49-50; 57pp; English.
 CC This protein comprises a novel rhesus monkey receptor, designated
 CC neuropeptide Y2 receptor, that has affinity for neuropeptide Y,
 CC pancreatic polypeptide and peptide YY. Y2 receptors are found in
 CC a variety of brain regions as well as in the peripheral nervous
 CC system. In all tissues, Y2 receptors mediate a decrease in the
 CC release of neurotransmitters. Also claimed are cDNA and RNA
 CC sequences (see AAV05530-31), expression vectors, host cells
 CC (especially Escherichia coli or mammalian cells), and methods for
 CC evaluating the effectiveness of test compounds for the treatment or
 CC prevention of conditions associated with an excess or deficiency of
 CC stimulation of the receptor. Such compounds can be used to treat or
 CC prevent disorders or diseases pertaining to the heart, blood vessels
 CC or the renal system (e.g. vasospasm, heart failure, shock, cardiac
 CC hypertrophy, increased blood pressure, angina, myocardial
 CC infarction, sudden cardiac death, arrhythmia, peripheral vascular
 CC disease and abnormal renal conditions such as impaired flow of
 CC fluid, abnormal mass transport or renal failure), conditions
 CC related to increased sympathetic nerve activity (e.g. during or
 CC after coronary artery surgery and operations and surgery in the
 CC gastrointestinal tract), cerebral diseases and diseases related to
 CC the central nervous system (e.g. epilepsy, cerebral infarction,
 CC neurodegeneration, stroke, cerebral vasospasm and haemorrhage,
 CC depression, anxiety, schizophrenia and dementia), conditions related
 CC to pain or nociception, diseases related to abnormal
 CC gastrointestinal motility and secretion (e.g. different forms of
 CC ileus, urinary incontinence, and Crohn's disease, abnormal drink
 CC and food intake disorders, such as obesity, anorexia, bulimia and
 CC metabolic disorders), diseases related to sexual dysfunction and
 CC reproductive disorders, conditions or disorders associated with
 CC inflammation, respiratory diseases (e.g. bronchoconstriction and
 CC asthma) and diseases related to abnormal hormone release (e.g.
 CC luteinizing hormone, growth hormone, insulin and prolactin).
 XX
 SQ Sequence 381 AA:

Matches	6;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	1	TRISKL	7						
Db	240	trwskl	246						

Search completed: July 15, 2002, 12:57:55
 Job time: 411 sec

Query Match 84.8%; Score 28; DB 19; Length 381;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:59:16 ; Search time 75.67 Seconds
(without alignments)
2.260 Million cell updates/sec

Title: US-09-712-819A-2

Perfect score: 33

Sequence: 1 TRIRSKL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	87.9	591	1	US-08-188-582-7	Sequence 7, Appl1
2	29	87.9	591	1	US-08-646-715-7	Sequence 7, Appl1
3	28	84.8	381	1	US-08-192-288-2	Sequence 2, Appl1
4	28	84.8	381	2	US-08-687-355A-2	Sequence 2, Appl1
5	28	84.8	381	2	US-08-687-355A-4	Sequence 6, Appl1
6	28	84.8	381	2	US-08-687-355A-6	Sequence 16, Appl1
7	28	84.8	381	4	US-09-200-673-16	Sequence 2, Appl1
8	28	84.8	381	4	US-08-876-798A-2	Sequence 4, Appl1
9	28	84.8	976	4	US-09-104-324B-4	Sequence 17, Appl1
10	28	84.8	584	1	US-08-313-288B-17	Sequence 33, Appl1
11	26	78.8	217	2	US-08-440-845D-3	Sequence 3, Appl1
12	26	78.8	411	2	US-08-868-458-3	Sequence 16, Appl1
13	26	78.8	411	4	US-09-522-800-16	Sequence 2, Appl1
14	26	78.8	476	4	US-08-924-183-1	Sequence 2, Appl1
15	26	78.8	476	4	US-08-924-183-2	Sequence 2, Appl1
16	26	78.8	476	4	US-08-924-183-7	Sequence 1, Appl1
17	26	78.8	476	4	US-09-488-364-1	Sequence 7, Appl1
18	26	78.8	476	4	US-09-488-364-2	Sequence 2, Appl1
19	26	78.8	476	4	US-09-488-364-7	Sequence 2, Appl1
20	26	78.8	479	1	US-08-278-635B-2	Sequence 2, Appl1
21	26	78.8	479	3	US-08-464-258B-2	Sequence 2, Appl1
22	26	78.8	479	3	US-08-471-961-1	Sequence 2, Appl1
23	26	78.8	484	4	US-08-913-578-2	Sequence 2, Appl1
24	26	78.8	484	4	US-08-785-427-2	Sequence 2, Appl1
25	26	78.8	729	2	US-08-677-298-2	Sequence 4, Appl1
26	26	78.8	946	3	US-08-560-005-4	Sequence 4, Appl1
27						

28	26	78.8	946	4	US-09-418-540-4	Sequence 4, Appl1
29	25	75.8	112	4	US-09-404-671-8	Sequence 8, Appl1
30	25	75.8	181	1	US-08-418-444A-3	Sequence 3, Appl1
31	25	75.8	218	4	US-09-651-656-105	Sequence 105, App
32	25	75.8	257	4	US-09-183-217-2	Sequence 2, Appl1
33	25	75.8	509	3	US-08-369-822C-26	Sequence 26, Appl1
34	25	75.8	509	3	US-08-582-776C-41	Sequence 41, Appl1
35	25	75.8	509	3	US-08-434-831B-38	Sequence 38, Appl1
36	24	72.7	213	1	US-08-287-442-2	Sequence 2, Appl1
37	24	72.7	213	1	US-08-459-701-2	Sequence 2, Appl1
38	24	72.7	213	1	US-08-460-298-2	Sequence 2, Appl1
39	24	72.7	213	1	US-08-459-174-2	Sequence 2, Appl1
40	24	72.7	213	1	US-08-761-258-3	Sequence 3, Appl1
41	24	72.7	213	1	US-08-761-258-5	Sequence 5, Appl1
42	24	72.7	213	1	US-08-761-258-10	Sequence 10, Appl1
43	24	72.7	213	2	US-09-009-218-2	Sequence 2, Appl1
44	24	72.7	213	2	US-09-009-218-4	Sequence 4, Appl1
45	24	72.7	213	2	US-09-009-218-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-188-582-7
; Sequence 7, Application US/08188582
; Patent No. 5534410
; GENERAL INFORMATION:
; APPLICANT: Tjian, Robert
; APPLICANT: Comai, Lucio
; APPLICANT: Dynlacht, Brian D.
; APPLICANT: Hoey, Timothy
; APPLICANT: Ruppert, Siegfried
; APPLICANT: Tanese, Naoko
; APPLICANT: Wang, Edith
; APPLICANT: Weinzierl, Robert O.J.
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TATS AND METHODS OF USE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,582
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-57650-2/AUT/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-188-582-7

Query Match 87.9%; Score 29; DB 1; Length 591;

Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0;

Indels 0;

Gaps 0;

QY 1 TRIFSK 6
|||||

Db 337 TRIFSK 342

RESULT 2
US-08-646-715-7

; Sequence 7, Application US/08646715

; Patent No. 5637686

; GENERAL INFORMATION:

; APPLICANT: Tjian, Robert

; APPLICANT: Comai, Lucio

; APPLICANT: Dylact, Brian D.

; APPLICANT: Hoey, Timothy

; APPLICANT: Ruppert, Siegfried

; APPLICANT: Tanese, Naoko

; APPLICANT: Wang, Edith

; APPLICANT: Weinzierl, Robert O.J.

; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS.

; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TATS AND METHODS OF USE

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/646,715

; FILING DATE: 09-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/188,582

; FILING DATE: 28-JAN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELEX: 910 277299

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 591 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-646-715-7

Query Match 87.9%; Score 29; DB 1; Length 591;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSK 6
|||||

Db 337 TRIFSK 342

RESULT 3

US-08-192-288-2

; Sequence 2, Application US/081922288

; Patent No. 5545549

; GENERAL INFORMATION:

; APPLICANT: Gerald, Christophe

; APPLICANT: Walker, Mary

; APPLICANT: Branchek, Theresa

; APPLICANT: Weinshank, Richard L.

; TITLE OF INVENTION: DNA ENCODING A HUMAN NEUROPEPTIDE

; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTOR AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/192,288

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 44742/JPW/TEP

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; TELEX: 422523 COOP UI

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 381 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-192-288-2

Query Match 84.8%; Score 28; DB 1; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TRIFSK 7
|||||

Db 240 TRIFSK 246

RESULT 4

US-08-687-355A-2

; Sequence 2, Application US/08687355A

; Patent No. 5989834

; GENERAL INFORMATION:

; APPLICANT: Synaptic Pharmaceutical Corporation

; TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE

; TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF

; NUMBER OF SEQUENCES: 27

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/687,355A
FILING DATE: No. 5989834ember 26, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPM/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-355A-2

Query Match 84.8%; Score 28; DB 2; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||:||||
DB 240 TRIMSKL 246

RESULT 5
US-08-687-355A-4
Sequence 4, Application US/08687355A
Patent No. 5989834
GENERAL INFORMATION:
APPLICANT: Synaptic Pharmaceutical Corporation
TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,355A
FILING DATE: No. 5989834ember 26, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPM/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-355A-4

Query Match 84.8%; Score 28; DB 2; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||:||||
DB 240 TRIMSKL 246

RESULT 6
US-08-687-355A-6
Sequence 6, Application US/08687355A
Patent No. 5989834
GENERAL INFORMATION:
APPLICANT: Synaptic Pharmaceutical Corporation
TITLE OF INVENTION: NUCLEIC ACID ENCODING NEUROPEPTIDE
TITLE OF INVENTION: Y/PEPTIDE YY (Y2) RECEPTORS AND USES THEREOF
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687,355A
FILING DATE: No. 5989834ember 26, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28, 678
REFERENCE/DOCKET NUMBER: 44742-A-PCT/JPM/MAT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-687-355A-6

Query Match 84.8%; Score 28; DB 2; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||:||||
DB 240 TRIMSKL 246

RESULT 7
US-09-200-673-16
Sequence 16, Application US/09200673A
Patent No. 6316203
GENERAL INFORMATION:
APPLICANT: Gerald, Christophe P.G.
APPLICANT: Weinschank, Richard L.
APPLICANT: Walker, Mary W.
APPLICANT: Branchek, Theresa
TITLE OF INVENTION: Methods of Modifying Feeding Behavior, Compounds Useful
TITLE OF INVENTION: In Such Methods, and DNA Encoding A Hypothalamic
FILE REFERENCE: 46166-B2/JPM
CURRENT APPLICATION NUMBER: US/09/200,673A
CURRENT FILING DATE: 1998-11-25
EARLIER APPLICATION NUMBER: 08/566,096
EARLIER FILING DATE: 1995-12-01
EARLIER APPLICATION NUMBER: 08/349,025

EARLIER FILING DATE: 1994-12-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 381
TYPE: PRT
ORGANISM: Homo sapiens
US-09-200-673-16

Query Match 84.8%; Score 28; DB 4; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSXL 7
|||:||||
Db 240 TRIMSXL 246

RESULT 8
US-08-876-798A-2
Sequence 2, Application US/08876798A
Patent No. 6355478
GENERAL INFORMATION:
APPLICANT: Baez, Melvyn
APPLICANT: Yang, Pelyi
TITLE OF INVENTION: RHESUS MONKEY NEUROPEPTIDE Y RECEPTOR
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/876,798A
FILING DATE: 16-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36, 808
REFERENCE/DOCKET NUMBER: X-10901
TELEPHONE: (317)276-0756
TELEFAX: (317)276-2763
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-876-798A-2

Query Match 84.8%; Score 28; DB 4; Length 381;
Best Local Similarity 85.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSXL 7
|||:||||
Db 240 TRIMSXL 246

RESULT 9
US-09-104-324B-4
Sequence 4, Application US/09104324B
Patent No. 6232460

GENERAL INFORMATION:
APPLICANT: T rec1, Ozlem; Sahin, ugru; Pfrendschuh, Michael
TITLE OF INVENTION: Methods for Diagnosis And Treating Cancers.
TITLE OF INVENTION: And Methods for Identifying Pathogenic Markers In A Sample
TITLE OF INVENTION: No. 6232460mat Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/104,324B
FILING DATE: 25-June-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/892,702
FILING DATE: 15-July-1997
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6232460man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5491
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3000
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 976 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-09-104-324B-4

Query Match 84.8%; Score 28; DB 4; Length 976;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSXL 7
|||:||||
Db 103 SRVFSXL 109

RESULT 10
US-08-313-288B-17
Sequence 17, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 584 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-17

Query Match 81.8%; Score 27; DB 1; Length 584;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||||:
DB 290 TRIFTKV 296

RESULT 11

US-08-297-431B-33
Sequence 33, Application US/08297431B
Patent No. 6136605
GENERAL INFORMATION:
APPLICANT: Fahl, William E
APPLICANT: Gulick, Andrew M
APPLICANT: Manoharan, T Herbert
APPLICANT: Puchalski, Ralph B
APPLICANT: Kramer, Katharine
APPLICANT: Wasserman Myeth W
TITLE OF INVENTION: NOVEL GLUTATHIONE S-TRANSFERASE ISOFORMS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saul, Ewing, Remick & Saul, LLP
STREET: 1500 Market Street, 38th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19102-2186
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/297,431B
FILING DATE: August 26, 1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Reed, Janet E.
REGISTRATION NUMBER: 36252
REFERENCE/DOCKET NUMBER: WARP F039
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-972-8386
TELEFAX: 215-972-2292
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 217 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-297-431B-33

Query Match 78.8%; Score 26; DB 4; Length 217;

Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||||
DB 205 TRIFSKL 211

RESULT 12

US-08-440-845D-3
Sequence 3, Application US/08440845D
Patent No. 5955329
GENERAL INFORMATION:
APPLICANT: Yuan, L.
APPLICANT: Kridl, J.
APPLICANT: Dehesh, K.
APPLICANT: Knaut, V.
TITLE OF INVENTION: Engineering Plant Thioesterases For
TITLE OF INVENTION: Altered Substrate Specificity.
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.40 MB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,845D
FILING DATE: 15-MAY-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schmedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 113
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 residues
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-845D-3

Query Match 78.8%; Score 26; DB 2; Length 411;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TRIFSKL 7
|||||
DB 259 TRIFSKL 265

RESULT 13

US-08-868-458-3
Sequence 3, Application US/08868458
Patent No. 6150512
GENERAL INFORMATION:
APPLICANT: Yuan, L.
TITLE OF INVENTION: Engineering Plant Thioesterases And
Disclosure of Plant Thioesterases
Having No. 6150512el Substrate Specificity
NUMBER OF SEQUENCES: 19

```

CORRESPONDENCE ADDRESS:
ADDRESS: Calgene, Inc.
STREET: 1920 Fifth street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: Microsoft Word For Window 95 7.0a

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,458
FILING DATE: 03-Jun-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/07064
FILING DATE: 15-MAY-96
APPLICATION NUMBER: 08/537,083
FILING DATE: 29-SEPT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: CGNE 133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 411 residues
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-08-868-458-3

Query Match      78.8% Score 26; DB 4; Length 411;
Best Local Similarity 85.7% Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRFSTKL 7
DB 259 TRFSTKL 265

RESULT 14
US-09-522-800-16
Sequence 16, Application US/09522800
Patent No. 6211164
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Ian, Yuo
APPLICANT: Giranda, Vincent L.
APPLICANT: Rockow-Magnone, Shayna K.
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES OF THE HUMAN
FILE REFERENCE: 6675, US 01
CURRENT APPLICATION NUMBER: US/09/522,800
CURRENT FILING DATE: 2000-03-10
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 476
TYPE: PRT
ORGANISM: Homo sapiens
US-09-522-800-16

Query Match      78.8% Score 26; DB 4; Length 476;
Best Local Similarity 71.4% Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 TRFSTKL 7
DB 378 TRFSTKL 384
```

```
RESULT 15
US-08-924-183-1
Sequence 1, Application US/08924183A
Patent No. 6218109
GENERAL INFORMATION:
APPLICANT: Elledge, Stephen J.
APPLICANT: Sanchez, Yolanda
TITLE OF INVENTION: MAMMALIAN CHECKPOINT GENES AND PROTEINS
FILE REFERENCE: 120541-1003
CURRENT APPLICATION NUMBER: US/08/924,183A
CURRENT FILING DATE: 1997-09-05
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 1
LENGTH: 476
TYPE: PRT
ORGANISM: Homo sapiens
US-08-924-183-1
```

```
Query Match      78.8% Score 26; DB 4; Length 476;
Best Local Similarity 71.4% Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TRFSTKL 7
DB 378 TRFSTKL 384
```

Search completed: July 15, 2002, 12:59:17
Job time: 388 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:00:59 ; Search time 95.45 Seconds
(without alignments)
7.047 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35

Sequence: 1 FYQLALT 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	88.6	628	2	S50524	hypothetical prote
2	30	85.7	129	2	S45868	probable membrane
3	30	85.7	317	2	T16710	hypothetical prote
4	30	85.7	857	2	E95009	cell wall surface
5	30	85.7	1161	2	C97881	conserved hypotet
6	29	82.9	211	2	H64216	ribosomal protein
7	29	82.9	350	2	B82230	conserved hypotet
8	29	82.9	380	2	T43016	DNA mismatch repai
9	29	82.9	658	2	E71465	hypothetical prote
10	29	82.9	658	2	G81727	conserved hypotet
11	29	82.9	684	2	T50317	probable DNA misma
12	29	82.9	769	2	S54525	mismatch repair pr
13	29	82.9	785	2	G75578	aculeacin A acylas
14	28	80.0	156	1	P69857	conserved hypotet
15	28	80.0	178	2	S44624	C50C3.5 protein -
16	28	80.0	179	2	B83903	RNA polymerase ECF
17	28	80.0	237	2	JS0594	rod-core linker po
18	28	80.0	237	2	AG1873	phycobillosome rod-
19	28	80.0	241	2	T25506	hypothetical prote
20	28	80.0	317	2	H95902	probable sugar ABC
21	28	80.0	480	2	S76598	hypothetical prote
22	28	80.0	483	2	H71482	probable ABC trans
23	28	80.0	483	2	G81745	conserved hypotet
24	28	80.0	501	2	T16084	hypothetical prote
25	28	80.0	541	1	MMBYN2	glucose transport
26	28	80.0	546	1	S48313	hexose transport p
27	28	80.0	566	2	S51081	glucose transport
28	28	80.0	668	2	T38362	hypothetical colle
29	28	80.0	699	2	S62408	hypothetical prote

30	28	80.0	783	2	AH2226	hypothetical prote
31	28	80.0	905	2	T23510	hypothetical prote
32	28	80.0	2130	2	AB0821	probable exported
33	27	77.1	89	2	I52695	cyclin-dependent k
34	27	77.1	105	2	D69619	multidrug resistant
35	27	77.1	153	2	S76676	hypothetical prote
36	27	77.1	221	2	S39522	hypothetical prote
37	27	77.1	240	1	Z1ZM91	19K zein precursor
38	27	77.1	240	1	Z1ZM2	19K zein precursor
39	27	77.1	240	1	Z1ZM92	19K zein precursor
40	27	77.1	240	2	T04364	19K zein precursor
41	27	77.1	240	2	A22831	19K zein precursor
42	27	77.1	249	2	E90048	glycerophosphodies
43	27	77.1	250	2	H83784	alpha kaifirin - so
44	27	77.1	267	2	T14824	ribonuclease S hom
45	27	77.1	269	2	T04095	

ALIGNMENTS

RESULT 1
S50524
hypothetical protein YEL065w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 23-Mar-2001
C:Accession: S50524
R:Dieterich, F.S.
Submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmid 9669, 8334, 8199, and lambda c10
A:Reference number: S50434
A:Accession: S50524
A:Molecule type: DNA
A:Residues: 1-628 <DIR>
A:Cross-references: EMBL:U18795; NID:9603241; PIDN:AAB65022.1; PID:9603253; GSPDB:GN0
A:Gene: MIPS:YEL065w
A:Map position: 5L
C:Superfamily: conserved probable membrane protein YEL073c

Query Match 88.6%; Score 31; DB 2; Length 628;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQLALT 7
DB 168 FYQLGIF 174

RESULT 2
S45868
probable membrane protein YBR013c - yeast (Saccharomyces cerevisiae)
M:Alternate names: hypothetical protein YBR0209
C:Species: Saccharomyces cerevisiae
C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 04-Mar-2000
C:Accession: S45868
R:Entian, K.D.; Koetter, P.; Rose, M.; Li, Z.; Therman, R.; Brendel, M.; Baur, A.; B
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45862
A:Accession: S45868
A:Molecule type: DNA
A:Residues: 1-119 <ENT>
A:Cross-references: EMBL:Z35882; NID:9577857; PID:9577860; GSPDB:GN00002; MIPS:YBR013
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YBR013c
A:Map position: 2R
C:Superfamily: Saccharomyces cerevisiae probable membrane protein YBR013c
C:Keywords: transmembrane protein
F:1-18/Domain: transmembrane #status predicted <TMM>

Query Match 85.7%; Score 30; DB 2; Length 129;
Best Local Similarity 85.7%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
|||||
Db 120 FYQLALT 126

RESULT 3

T16710

hypothetical protein R10H1.2 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16710

R:Bentley, D.

submitted to the EMBL Data Library, July 1995

A:Description: The sequence of *C. elegans* cosmid R10H1.

A:Reference number: Z18563

A:Accession: T16710

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-317 <BEN>

A:Cross-references: EMBL:U23174; NID:9726398; PID:g726400; PIDN:AAC46710.1; CESP:R10H1.2

C:Genetics:

A:Gene: CESP:R10H1.2

A:Introns: 12/3; 65/3; 126/2; 182/3; 251/1; 279/3

Query Match

Best Local Similarity 85.7%; Score 30; DB 2; Length 317;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
|||||
Db 124 FYQLALT 130

RESULT 4

E95009

cell wall surface anchor family protein [Imported] - *Streptococcus pneumoniae* (strain T16710)

C:Species: *Streptococcus pneumoniae*

C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001

C:Accession: E95009

R:Rettelin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfe,

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A>Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: E95009

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-857 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK74270.1; PID:g14971548; GSPDB:GN00164; TIGR:SP4

C:Genetics:

A:Experimental source: strain TIGR4

A:Gene: SP0082

Query Match

Best Local Similarity 85.7%; Score 30; DB 2; Length 857;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQLALT 6
|||||
Db 538 FYQLALT 543

RESULT 5..

C97881

conserved hypothetical protein spr0075 [Imported] - *Streptococcus pneumoniae* (strain C:Species: *Streptococcus pneumoniae*

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001

C:Accession: C97881

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.

e, R.; LeBlanc, D.J.; Lee, L.N.; Leikowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.

A>Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: C97881

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1161 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK98879.1; PID:g15457609; GSPDB:GN00174

C:Genetics:

A:Gene: spr0075

Query Match 85.7%; Score 30; DB 2; Length 1161;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQLALT 6
|||||
Db 842 FYQLALT 847

RESULT 6

H64216

ribosomal protein L4 - *Mycoplasma genitalium*

C:Species: *Mycoplasma genitalium*

C>Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 07-Dec-1999

C:Accession: H64216

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.

M.; Fuhmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick,

, C.A.; Venter, J.C.

Science 270, 397-403, 1995

A>Title: The minimal gene complement of *Mycoplasma genitalium*.

A:Reference number: A64200; MUID:96026346

A:Accession: H64216

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-211 <TIGR>

A:Cross-references: GB:U39695; GB:I43967; NID:g1045833; PID:g1045836; TIGR:MG152

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: SCC3

C:Superfamily: Escherichia coli ribosomal protein L4

Query Match

Best Local Similarity 82.9%; Score 29; DB 2; Length 211;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YQLALT 7
|||||
Db 107 YQLALT 112

RESULT 7

B82230

conserved hypothetical protein VC1194 [Imported] - *Vibrio cholerae* (strain N16961 ser

C:Species: *Vibrio cholerae*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82230

R:Heidelberg, J.F.; Eisen, J.A.; Neilson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A:Reference number: AB2035; MUID:20406833

A:Accession: B82230
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-350 <HEI>
A:Cross-references: GB:AE004199; GB:AE003852; NID:g9655666; PIDN:AAF94353.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype EI Tor
C:Genetics:
A:Gene: VC1194
A:Map position: 1

Query Match 82.9%; Score 29; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FYOLALT 7
|||||
DB 50 FYOLALT 55

RESULT 8
T43016
DNA mismatch repair protein MLH1 homolog - fission yeast (Schizosaccharomyces pombe) (fr
C:Species: Schizosaccharomyces pombe
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43016
R:Oshiohoka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A:Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: Z17323; MUID:98162722
A:Accession: T43016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-380 <TOS>
A:Cross-references: EMBL:D89212; NID:g1749631; PIDN:BA13873.1; PID:g1749632
A:Experimental source: strain PR745
C:Superfamily: mismatch repair protein MLH1

Query Match 82.9%; Score 29; DB 2; Length 380;
Best Local Similarity 71.4%; Pred. No. 46;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||:|
DB 203 FYOLALT 209

RESULT 9
E71465
hypothetical protein CT837 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 18-Aug-2000
C:Accession: E71465
R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A:Reference number: A71570; MUID:99000809
A:Accession: E71465
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-658 <ARN>
A:Cross-references: GB:AE001356; GB:AE001273; NID:g3329299; PIDN:AAC68434.1; PID:g332930
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: CT837
C:Superfamily: conserved hypothetical protein TC0225

Query Match 82.9%; Score 29; DB 2; Length 658;
Best Local Similarity 85.7%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
||| |||
DB 217 FYOLALT 223

RESULT 10
G81727
conserved hypothetical protein TC0225 [imported] - Chlamydia muridarum (strain N199)
C:Species: Chlamydia muridarum Chlamydia trachomatis MoPn
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000
C:Accession: G81727
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: G81727
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-658 <TEP>
A:Cross-references: GB:AF002289; GB:AF002160; NID:g7190254; PIDN:AAF39097.1; PID:g719
A:Experimental source: strain N199 (MoPn)
C:Genetics:
A:Gene: TC0225
C:Superfamily: conserved hypothetical protein TC0225

Query Match 82.9%; Score 29; DB 2; Length 658;
Best Local Similarity 85.7%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
||| |||
DB 217 FYOLALT 223

RESULT 11
T50317
probable DNA mismatch repair protein, MLH1 homolog [imported] - fission yeast (Schizo
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000
C:Accession: T50317
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; LeJaure, V.; Galibert
submitted to the EMBL Data Library, January 2000
A:Reference number: Z25061
A:Accession: T50317
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-684 <MCD>
A:Cross-references: EMBL:AL136536; PIDN:CAB66448.1; GSPDB:GN00067; SPDB:SPBC1703.04
A:Experimental source: strain 972h(-); cosmid c1703
C:Genetics:
A:Gene: SPDB:SPBC1703.04
A:Map position: 2
A:Introns: 24/3; 70/3; 128/2
C:Superfamily: mismatch repair protein MLH1

Query Match 82.9%; Score 29; DB 2; Length 684;
Best Local Similarity 71.4%; Pred. No. 86;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||:|
DB 507 FYOLALT 513

RESULT 12
S54525
mismatch repair protein MLH1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YH8520.16; protein YMR167W
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Sep-1999

C:Accession: S54525; S48890; S54612
R: Hunt, S.; Bowman, S.
Submitted to the EMBL Data Library, May 1995
A: Reference number: S54510
A: Accession: S54525
A: Molecule type: DNA
A: Residues: 1-769 <HUN>
A: Cross-references: GB:249705; EMBL:249700; NID:9825556; PIDN:CAA89803.1; PID:9825572; E
A: Experimental source: strain AB972
R: Polla, T.A.; Christie, D.M.; Liskay, R.M.
Mol. Cell. Biol. 14, 407-415, 1994
A: Title: Dual requirement in yeast DNA mismatch repair for MLH1 and PMS1, two homologs
A: Reference number: S48890; MUID:94088538
A: Accession: S48890
A: Molecule type: DNA
A: Residues: 1-257, 'L', 259-287, 'F', 289-707, 'L', 709-769 <PRO>
A: Cross-references: EMBL:U07187; NID:9460626; PIDN:AAA16835.1; PID:9460627
C: Genetics:
A: Gene: SCD:MLH1
A: Cross-references: SCD:SO004777; MIPS:YMR167w
A: Map position: 13R
C: Function:
A: Description: DNA repair
A: Superfamily: mismatch repair protein MLH1
C: Keywords: DNA repair

Query Match 82.9%; Score 29; DB 2; Length 769;
Best Local Similarity 71.4%; Pred. No. 98;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||: ||
DB 570 FYQIGLT 576

RESULT 13
G75578
Baculacin A acylase - *Deinococcus radiodurans* (strain R1)
C: Species: *Deinococcus radiodurans*
C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C: Accession: G75578
R: White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A: Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A: Reference number: A75250; MUID:20036896
A: Accession: G75578
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-785 <WHI>
A: Cross-references: GB:AE001863; GB:AE001825; NID:96460670; PIDN:AAF12385.1; PID:9646068
A: Experimental source: strain R1
C: Genetics:
A: Gene: DRA0255
A: Map position: 2

Query Match 82.9%; Score 29; DB 2; Length 785;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 FYOLALT 7
|||||
DB 257 FYQLHLT 263

RESULT 14
F69857
conserved hypothetical protein ykna - *Bacillus subtilis*
C: Species: *Bacillus subtilis*
C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

C: Accession: F69857
R: Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A: Authors: Foulger, D.; Fritz, C.; Fujita, M.; Funai, S.; Galizzi, A.; Gal
leoh, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A: Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
A: Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchiya
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A: Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A: Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A: Reference number: A69580; MUID:98044033
A: Accession: F69857
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-156 <KUN>
A: Cross-references: GB:299110; GB:AL009126; NID:92633472; PIDN:CAB13174.1; PID:926336
A: Experimental source: strain 168
C: Genetics:
A: Gene: ykna
C: Superfamily: hypothetical protein yaaJ

Query Match 80.0%; Score 28; DB 1; Length 156;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 FYOLALT 7
|||: ||
DB 132 FYQVTLT 138

RESULT 15
S44624
C50C3.5 protein - *Caenorhabditis elegans*
C: Species: *Caenorhabditis elegans*
C: Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 30-Jun-2001
C: Accession: S44624
R: Du, Z.
submitted to the EMBL Data Library, May 1993
A: Description: Sequence of the C. elegans cosmid C50C3.
A: Reference number: S44626
A: Accession: S44624
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-178 <FAV>
A: Cross-references: EMBL:L14433; NID:9289649; PIDN:AAA27976.1; PID:9289657
C: Genetics:
A: Introns: 51/3; 117/3; 144/2; 162/1
C: Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology
F: 47-79/Domain: calmodulin repeat homology <EF1>
F: 83-115/Domain: calmodulin repeat homology <EF2>

Query Match 80.0%; Score 28; DB 2; Length 178;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 FYOLALT 7
:|:||||
DB 160 YRLALT 166

Search completed: July 15, 2002, 13:01:01
Job time: 467 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:23:31 ; Search time 44.9 Seconds
(without alignments)
6.036 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35
Sequence: 1 FYQLALF 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	88.6	628	YEHS_YEAST	P39980 saccharomyc
2	30	85.7	129	YBM3_YEAST	P38215 saccharomyc
3	30	85.7	317	YRP2_CAEEL	Q09344 caenorhabdl
4	29	82.9	111	NIFW_SYNP8	Q07360 synchococc
5	29	82.9	211	RL4_MYC6E	P47398 mycoplasma
6	29	82.9	769	MLH1_YEAST	P38920 saccharomyc
7	28	80.0	178	YLJ5_CAEEL	P34368 caenorhabdl
8	28	80.0	236	PYG3_ANASP	P29888 anabaena sp
9	28	80.0	480	Y074_SYNP3	Q55790 synchocyst
10	28	80.0	541	HXR2_YEAST	P23585 saccharomyc
11	28	80.0	546	HXR2_YEAST	P43581 saccharomyc
12	28	80.0	566	KHRT_KIULA	P53387 kluyveromyc
13	28	80.0	668	YAT6_SCHPO	O09761 schizosacch
14	27	77.1	105	EBRA_BACSU	O31792 bacillus su
15	27	77.1	221	YC24_GALSU	P35912 galidieria s
16	27	77.1	240	ZEAB_MAIZE	P04702 zea mays (m
17	27	77.1	240	ZEAB_MAIZE	P04703 zea mays (m
18	27	77.1	240	ZEAB_MAIZE	P06676 zea mays (m
19	27	77.1	240	ZEAB_MAIZE	P06677 zea mays (m
20	27	77.1	279	YJ86_YEAST	P47153 saccharomyc
21	27	77.1	298	PK1_ASFB7	P42493 african swi
22	27	77.1	299	PK1_ASFB2	P34206 african swi
23	27	77.1	303	CDK4_HUMAN	P11802 homo sapien
24	27	77.1	303	CDK4_MOUSE	P30285 mus musculu
25	27	77.1	303	CDK4_PIG	P79432 sus scrofa
26	27	77.1	303	CDK4_RAT	P35426 ratus norv
27	27	77.1	328	THIL_HAEIN	Q57190 haemophilus
28	27	77.1	354	E2B_XYLFA	Q99495 xylella fas
29	27	77.1	483	EXON_HSVSA	Q01013 herpesvirus
30	27	77.1	483	YC24_GUTTH	Q78473 guillardia
31	27	77.1	485	ENGA_NEIGO	O87407 neisseria g
32	27	77.1	486	YC24_CYAPA	P48260 cyanophora
33	27	77.1	486	YC24_ODOSI	P49530 odontella s

34	27	77.1	487	1	YC24_PORPU	P51240 porphyra pu
35	27	77.1	495	1	SUFB_ECOLI	P77522 escherichia
36	27	77.1	618	1	SL55_RAT	O63008 ratus norv
37	27	77.1	643	1	SL55_HUMAN	O92911 homo sapien
38	27	77.1	762	1	ACCX_BOVIN	O29437 bos taurus
39	27	77.1	765	1	PFLD_ECOLI	P32674 escherichia
40	27	77.1	949	1	SRB4_YEAST	P50104 saccharomyc
41	27	77.1	1376	1	CRBH_HUMAN	P82279 homo sapien
42	27	77.1	1638	1	DPO3_LACIA	Q9cdt7 lactococcus
43	27	77.1	2039	1	CCH1_YEAST	P50077 saccharomyc
44	26	74.3	72	1	YAA1_ECOLI	P75620 escherichia
45	26	74.3	91	1	YVBE_VACCV	P20545 vaccinia vi

ALIGNMENTS

RESULT	ID	YEHS_YEAST	STANDARD	PRT	628 AA.
AC	P39980:	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)				
DT	01-FEB-1995 (Rel. 31, Last annotation update)				
DE	Hypothetical 70.6 kDa protein in HXR2-CANI intergenic region.				
DN	YEL065W.				
OS	Saccharomyces cerevisiae (Baker's yeast).				
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.				
OX	NCBI_TaxID=4932;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-S288C / AB972;				
RA	Dierich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,				
RA	Aviles E., Berio A., Brennan T., Carpenter J., Chen E., Cherry J.M.,				
RA	Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,				
RA	Hymen R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,				
RA	Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,				
RA	Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,				
RA	Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;				
RL	Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
DR	EMBL; 018795; AAB65022.1; .				
DR	SGD; S0000791; SIR1.				
KW	Hypothetical protein; Transmembrane.				
FT	TRANSMEM 68				
FT	TRANSMEM 107				
FT	TRANSMEM 127				
FT	TRANSMEM 132				
FT	TRANSMEM 164				
FT	TRANSMEM 184				
FT	TRANSMEM 194				
FT	TRANSMEM 225				
FT	TRANSMEM 285				
FT	TRANSMEM 317				
FT	TRANSMEM 337				
FT	TRANSMEM 354				
FT	TRANSMEM 374				
FT	TRANSMEM 394				
FT	TRANSMEM 420				
FT	TRANSMEM 448				
FT	TRANSMEM 488				
FT	TRANSMEM 508				
FT	TRANSMEM 559				
SO	SEQUENCE 628 AA; 70561 MW; 39E711CF861CF05 CRC64;				

Query Match Score 31; DB 1; Length 628;
Best Local Similarity 85.7%; Pred. No. 12;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
 |||||
 Db 168 FYQALGT 174

RESULT 2
 YBM3_YEAST STANDARD; PRT; 129 AA.
 ID YBM3_YEAST
 AC P38215;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Hypothetical 14.5 kDa protein in IPPI-TTPI intergenic region precursor.
 GN YBR013C OR YBR0209.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=8288C;
 RA Entian K.-D., Koetter P., Rose M., Li Z., Therman R., Brendel M.,
 RA Baur A., Boles E., Miosga T., Schaaf-Gerstenschlaeger I.,
 RA Zimmermann F.K.;
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 DR EMBL; 235882; CAAB4955.1; -
 DR PIR; S45868; S45868.
 DR SGD; S0000217; YBR013C.
 DR KW Hypothetical protein; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 1 129 HYPOTHELICAL PROTEIN YBR013C.
 FT SEQUENCE 129 AA; 14481 MW; DBE90D47E09644E9 CRC64;
 SQ

Query Match 85.7%; Score 30; DB 1; Length 129;
 Best Local Similarity 85.7%; Pred. No. 3.9;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
 |||||
 Db 120 FYQALGT 126

RESULT 3
 YRP2_CABEL STANDARD; PRT; 317 AA.
 ID YRP2_CABEL
 AC Q09344;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 35.8 kDa protein R10H1.2 from chromosome II.
 GN R10H1.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL NZ;
 RA Bentley D.;
 RL Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: SOME, TO C. ELEGANS F27E5.5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----

DR EMBL; 023174; AAC6710.1; -
 DR WormPep; R10H1.2; CE02029.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR004151; Strc.
 DR Pfam; PF03125; Strc; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECIP_F1.2; UNKNOWN_1.
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 105 125 POTENTIAL.
 FT TRANSMEM 147 167 POTENTIAL.
 FT TRANSMEM 191 211 POTENTIAL.
 FT TRANSMEM 232 252 POTENTIAL.
 FT TRANSMEM 292 312 POTENTIAL.
 SQ SEQUENCE 317 AA; 35770 MW; 88B9BB32EF15D4E2 CRC64;

Query Match 85.7%; Score 30; DB 1; Length 317;
 Best Local Similarity 85.7%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
 |||||
 Db 124 FYHALAT 130

RESULT 4
 NIFW_SYNP8 STANDARD; PRT; 111 AA.
 ID NIFW_SYNP8
 AC O07360;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Nitrogenase stabilizing/protective protein nifW.
 GN NIFW.
 OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
 OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
 OX NCBI_TaxID=41431;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99231861; PubMed-10217509;
 RA Huang T.-C., Lin R.-F., Chu M.-K., Chen H.-M.;
 RT "Organization and expression of nitrogen-fixation genes in the aerobic
 RT nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain
 RF-1."
 RL Microbiology 145:743-753(1999).
 CC -1- FUNCTION: May protect the nitrogenase Fe-Mo protein from oxidative
 CC damage (By similarity).
 CC -1- SUBUNIT: Homotrimer; associates with nifD (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE NIFW FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 DR EMBL; AF003700; AAC35197.1; -
 DR NITROGEN fixation.
 KW NITROGEN fixation.
 SQ SEQUENCE 111 AA; 12931 MW; C85B39697EFC310E CRC64;

Query Match 82.9%; Score 29; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YQALAT 7
 111111
 Db 67 YQALAT 72

RESULT 5

RL MYCGE STANDARD; PRT; 211 AA.
 AC P47398;
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 50S ribosomal protein L4.
 GN RPLD OR RPL4 OR MG152.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;

RA FRASER C.M., GOCAYNE J.D., WHITE O., ADAMS M.D., CLAYTON R.A.,
 RA FLEISCHMAN R.D., BULT C.J., KEILAVAGE A.R., SUTTON G., KELLEY J.M.,
 RA FRITZMAN J.L., WEIDMAN J.F., SMALL K.V., SANDUSKY M., FUHRMAN J.L.,
 RA NGUYEN D.T., UTTERBACK T.R., SAUDEK D.M., PHILLIPS C.A., MERRICK J.M.,
 RA TOMB J.F., DOUGHERTY B.A., BOLT R.F., HU P.-C., LUCIER T.S.,
 RA PETERSON S.N., SMITH H.O., HUTCHISON C.A. III, VENTER J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium.";
 RL Science 270:397-403(1995).

CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY AND SPECIFICALLY TO 23S
 CC RRNA (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE L4P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; U39696; AAC71370.1; -.
 DR TIGR; MG152; -.
 DR InterPro: IPR002136; Ribosomal_L4/L1E.
 DR Pfam: PF00573; Ribosomal_L4; 1.
 KW Ribosomal protein; rRNA-binding; Complete proteome.
 SO SEQUENCE 211 AA; 23508 MW; 8791B4689854D23 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YQALAT 7
 111111
 Db 107 YQALAT 112

RESULT 6
 MLH1_YEAST STANDARD; PRT; 769 AA.

AC P38920;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MOTL protein homolog 1 (DNA mismatch repair protein MLH1).
 GN MLH1 OR YMR167W OR YM6520.16.

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;

RA PROLLA T.A., CHRISTIE D.-M., LISKAY R.M.;
 RT "Dual requirement in yeast DNA mismatch repair for MLH1 and PMS1, two
 RT homologs of the bacterial mutL gene.";
 RL Mol. Cell. Biol. 14:407-415(1994).

CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
 CC IN DNA. PLAYS A MAJOR ROLE IN MAINTAINING THE GENETIC STABILITY
 CC OF SIMPLE SEQUENCE REPEATS AND IN THE REPAIR OF HETERO-DUPLEX
 CC SITES PRESENT IN MEIOTIC RECOMBINATION INTERMEDIATES. PROBABLY
 CC MSH2, PMS1, AND MLH1 ARE COMPONENTS OF THE SAME DNA MISMATCH
 CC REPAIR PATHWAY.

CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MOTL/HEXA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial/
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; U07187; AAA16835.1; -.
 DR EMBL; Z49705; CAA89803.1; -.
 DR PIR; S48890; S48890.
 DR HSP; P23367; IERN.
 DR SGD; S0004777; MLH1.
 DR InterPro: IPR002099; DNA_mis_repair.
 DR Pfam: PF01119; DNA_mis_repair; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
 DR DNA repair.
 KW CONFLICT 258 P -> L (IN REF. 1).
 FT CONFLICT 288 N -> F (IN REF. 1).
 FT CONFLICT 708 S -> L (IN REF. 1).
 SO SEQUENCE 769 AA; 87062 MW; B2DBB31DE3943171 CRC64;

Query Match 82.9%; Score 29; DB 1; Length 769;
 Best Local Similarity 71.4%; Pred. No. 45;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQALAT 7
 111111
 Db 570 FYQALAT 576

RESULT 7
 YLJ5_CAEEL STANDARD; PRT; 178 AA.

AC P34368;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hypothetical calcium-binding protein C50C3.5 in chromosome III.
 GN C50C3.5.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RC STRAIN-BRISTOL N2;
RA MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rikken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sullivan J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Welnstock L., Wilkinson-Sproat J.,
RA Wohlsman P.,
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.*
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: STRONG, TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L14433; AAA27976.1; -
DR PIR: S44624; S44624.
DR HSSP: P02593; 1CMG.
DR WormPep: C50C3.5; CE00121.
DR InterPro: IPR002068; EF-hand.
DR Pfam: PF00036; efhand; 2.
DR SMART: SM00054; EFh; 2.
DR PROSITE: PS00018; EF_HAND; 2.
DR Hypothetical protein; Calcium-binding; Repeat.
KW CA_BIND 60 71 EF_HAND 1 (POTENTIAL).
FT CA_BIND 96 107 EF_HAND 2 (POTENTIAL).
SQ SEQUENCE 178 AA; 20022 MW; EBRFC4C74D8C57BD CRC64;

Query Match
Best Local Similarity 71.4%; Score 28; DB 1; Length 178;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQLALT 7
Db 160 YRRLALT 166

RESULT 8
PYG3_ANASP STANDARD; PRT; 236 AA.
ID PYG3_ANASP
AC P29988;
DT 01-APR-1993 (Rel. 25; Created)
DT 01-OCT-1996 (Rel. 34; Last sequence update)
DT 01-MAR-2002 (Rel. 41; Last annotation update)
DE Phycobillosome rod-core linker polypeptide pPG3 (L-RC 27.2).
GN CPG3 OR ALR0536.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92077441; PubMed=1743523;
RA Bryant D.A., Stirewalt V.L., Glauser M., Frank G., Siedler W.,
RA Zuber H.;
RT "A small multigene family encodes the rod-core linker polypeptides of
RT Anabaena sp. PCC7120 phycobillosomes.";
RL Gene 107:91-99(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,

```

```

RA Watanabe A., Iriuchih M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shampo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.*;
RL DNA Res. 8:205-213(2001).
CC -1- FUNCTION: ROD-CORE LINKER PROTEIN REQUIRED FOR ATTACHMENT OF
CC PHYCOCYANIN TO ALLOPHYCOCYANIN IN CORES OF PHYCOBILISOMES.
CC AND THE LOCATION OF THE DISC-SHAPE PHYCOBILIPROTEIN UNITS WITHIN
CC THE PHYCOBILISOME AND MODULATE THEIR SPECTROSCOPIC PROPERTIES IN
CC ORDER TO MEDIANE A DIRECTED AND OPTIMAL ENERGY TRANSFER.
CC -1- SUBUNIT: THE PHYCOBILISOME IS A HEMIDISOCAL STRUCTURE THAT IS
CC COMPOSED OF TWO DISTINCT SUBSTRUCTURES: A CORE COMPLEX (THAT
CC CONTAINS PHYCOBILIPROTEINS) AND A NUMBER OF RODS RADIATING FROM
CC THE CORE.
CC -----
CC -1- SIMILARITY: TO OTHER PHYCOBILISOME LINKER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80435; AAA22038.1; -
DR PIR: AP003582; BAB72494.1; -
DR DR J50594; J50594.
DR InterPro: IPR001297; PBS_linker-poly.
DR Pfam: PF00427; PBS_linker-poly; 1.
KW Phycobillosome; Photosynthesis; Multigene family; Complete proteome.
FT INIT MET 0
SQ SEQUENCE 236 AA; 27074 MW; 54293B7074F147B9 CRC64;

Query Match
Best Local Similarity 71.4%; Score 28; DB 1; Length 236;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQLALT 7
Db 93 FYQLVYT 99

RESULT 9
Y074_SYNY3 STANDARD; PRT; 480 AA.
ID Y074_SYNY3
AC Q55790;
DT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical 52.8 kDa protein SLR0074.
GN SLR0074.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96127529; PubMed=8590279.
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.*;
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO THE UPF0051 (YCF24) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D64004; BAA10542.1; -;
 DR InterPro; IPR000825; UPP0051.
 DR Pfam; PF01458; UPP0051; 1.
 KW Hypothetical protein: Complete proteome.
 KW SEQUENCE 480 AA; 52795 MW; 066845CB17A0F33A CRC64;

Query Match 80.0%; Score 28; DB 1; Length 480;
 Best Local Similarity 71.4%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FTYOALIT 7
 11 : 111
 Db 328 FYSIALT 334

RESULT 10
 ID HXT2_YEAST STANDARD; PRT; 541 AA.
 AC P23585;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE High-affinity glucose transporter HXT2.
 GN HXT2 OR YMR011W OR YMR270.15.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

OX NCBI_Taxid-4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91042520; PubMed-2233722;
 RA Kruckeberg A.L., Bisson L.F.;
 RT "The HXT2 gene of Saccharomyces cerevisiae is required for
 RT high-affinity glucose transport";
 RL Mol. Cell. Biol. 10:5903-5913(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;

RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: HIGH-AFFINITY GLUCOSE TRANSPORTER. IS ONLY INDISPENSABLE
 CC FOR GROWTH ON LOW GLUCOSE-CONTAINING MEDIA, BECAUSE S.CEREVISIAE
 CC POSSESSES OTHER SUGAR TRANSPORTERS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- INDUCTION: RERESSED AT HIGH GLUCOSE CONCENTRATIONS.
 CC -1- MISCELLANEOUS: GLUCOSE TRANSPORT IS THOUGHT TO BE MEDIATED BY TWO
 CC KINETICALLY DISTINCT SYSTEMS, A GLUCOSE-REPRESSIBLE HIGH-AFFINITY
 CC SYSTEM AND A CONSTITUTIVE LOW-AFFINITY SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M33270; AAA34701.1; -;
 DR EMBL; Z48613; CAA86528.1; -;
 DR PIR; S12200; MMBYH2.
 DR SGD; S0004613; HXT2.
 DR InterPro; IPR003663; Sugar transporter.
 DR InterPro; IPR003662; sub transporter.
 DR Pfam; PF000683; sugar tr; 1
 DR PRINTS; PR00171; SUGRTNSPOT.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

KW Repeat: Transmembrane; Sugar transport; Transport; Glycoprotein;
 KW Phosphorylation.
 FT DOMAIN 1 49
 FT TRANSMEM 50 70
 FT DOMAIN 71 107
 FT TRANSMEM 108 128
 FT DOMAIN 129 134
 FT TRANSMEM 135 155
 FT DOMAIN 156 162
 FT DOMAIN 163 183
 FT TRANSMEM 184 192
 FT TRANSMEM 193 213
 FT DOMAIN 214 227
 FT TRANSMEM 228 248
 FT DOMAIN 249 327
 FT TRANSMEM 328 347
 FT DOMAIN 348 350
 FT TRANSMEM 351 371
 FT DOMAIN 372 379
 FT TRANSMEM 380 400
 FT DOMAIN 401 418
 FT TRANSMEM 419 439
 FT DOMAIN 440 456
 FT TRANSMEM 457 477
 FT DOMAIN 478 484
 FT TRANSMEM 485 505
 FT DOMAIN 506 541
 FT MOD_RES 266 266
 FT MOD_RES 539 539
 FT CARBOHYD 82 82
 SQ SEQUENCE 541 AA; 59840 MW; 6A6FEC0AB7391CA7 CRC64;

Query Match 80.0%; Score 28; DB 1; Length 541;
 Best Local Similarity 71.4%; Pred. No. 54;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FTYOALIT 7
 111 : 1
 Db 198 FTYOALMIT 204

RESULT 11
 ID HXT0_YEAST STANDARD; PRT; 546 AA.
 AC P43581;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Hexose transporter HXT10.
 GN HXT10 OR YFL011W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

OX NCBI_Taxid-4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RX MEDLINE-95400292; PubMed-7670463;
 RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
 RA Sasunuma S.-I., Sasunuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
 RA Yamazaki M., Tashiro H., Eki T.;
 RT "Analysis of the nucleotide sequence of chromosome VI from
 RT Saccharomyces cerevisiae";
 RL Nat. Genet. 10:261-268(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RA Barrell B.G., Churcher C., Rajandream M.A.;

RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D50617; BAA09227.1; -
DR EMBL: D31600; BAA06493.1; -
DR EMBL: Z46255; CAA86344.1; -
DR SGD: S0001883; HXT10.
DR InterPro: IPR003663; Sugar_transporter.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Repeat: Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 1 44
FT TRANSMEM 45 65
FT DOMAIN 66 100
FT TRANSMEM 101 121
FT TRANSMEM 122 127
FT TRANSMEM 128 158
FT TRANSMEM 149 179
FT TRANSMEM 180 185
FT TRANSMEM 186 206
FT TRANSMEM 207 220
FT TRANSMEM 221 241
FT TRANSMEM 242 324
FT TRANSMEM 325 341
FT TRANSMEM 342 347
FT TRANSMEM 348 365
FT TRANSMEM 366 372
FT TRANSMEM 373 393
FT TRANSMEM 394 415
FT TRANSMEM 416 436
FT TRANSMEM 437 453
FT TRANSMEM 454 474
FT TRANSMEM 475 475
FT TRANSMEM 476 496
FT TRANSMEM 497 546
FT CARBOHYD 75 75
SQ SEQUENCE 546 AA; 60662 MW; B2D80CDBBCC398F CRC64;

Query Match 80.0%; Score 28; DB 1; Length 546;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
DB 191 FYOLMIT 197

RESULT 12
KHT2_KLULA
ID KHT2_KLULA STANDARD; PRT; 566 AA.
AC P53387;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hexose transporter 2.
GN KHT2.
OS Kluyveromyces lactis (yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-JA6;
RX MEDLINE-98028406; PubMed-9363776;
RA Melich J, Goffini P, Kuger P, Ferrero I, Braunig K.D.;
RT "Influence of mutations in hexose transporter genes on glucose
RT repression in Kluyveromyces lactis.";
RL Eur. J. Biochem. 249:248-257(1997).
CC -1- FUNCTION: PROBABLE GLUCOSE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: Z47080; CAA87389.1; -
DR InterPro: IPR003663; Sugar_transporter.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PRINTS: PR00171; SUGRTRNSPORT.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
KW Repeat: Transmembrane; Sugar transport; Glycoprotein.
FT DOMAIN 22 61
FT TRANSMEM 62 82
FT TRANSMEM 83 112
FT TRANSMEM 113 133
FT TRANSMEM 134 140
FT TRANSMEM 141 161
FT TRANSMEM 162 166
FT TRANSMEM 167 187
FT TRANSMEM 188 198
FT TRANSMEM 199 219
FT TRANSMEM 220 233
FT TRANSMEM 234 254
FT TRANSMEM 255 333
FT TRANSMEM 334 353
FT TRANSMEM 354 357
FT TRANSMEM 358 378
FT TRANSMEM 379 385
FT TRANSMEM 386 406
FT TRANSMEM 407 428
FT TRANSMEM 429 449
FT TRANSMEM 450 465
FT TRANSMEM 466 486
FT TRANSMEM 487 492
FT TRANSMEM 493 513
FT TRANSMEM 514 566
FT CARBOHYD 88 88
SQ SEQUENCE 566 AA; 62727 MW; 9784173EC2375444 CRC64;

Query Match 80.0%; Score 28; DB 1; Length 566;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
DB 204 FYOLMIT 210

RESULT 13
YA76_SCHPO
ID YA76_SCHPO STANDARD; PRT; 668 AA.
AC Q09761;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 75.9 kDa protein C24H6.06 in chromosome I.
GN SPAC24H6.06.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:22:40 ; Search time 172.49 Seconds
(without alignments)
7.020 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35

Sequence: 1 FYQLALT 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp rodent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.rvirus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	91.4	1346	16 Q98IR5	Q98IR5 rhizobium 1
2	30	85.7	370	2 O52357	O52357 bacillus pu
3	30	85.7	744	8 Q9TL93	Q9TL93 villaria c
4	30	85.7	857	16 Q97T70	Q97T70 streptococ
5	29	82.9	111	2 O07360	O07360 synchococc
6	29	82.9	350	16 Q9KSR2	Q9KSR2 vibrio chol
7	29	82.9	380	3 P78862	P78862 schizosacch
8	29	82.9	658	16 O84844	O84844 chlamydia t
9	29	82.9	658	16 Q9PL82	Q9PL82 chlamydia m
10	29	82.9	684	3 Q9P7W6	Q9P7W6 schizosacch
11	29	82.9	755	5 Q9BKJ9	Q9BKJ9 dictyostell
12	29	82.9	785	16 Q9RYC4	Q9RYC4 deinococcus
13	28	80.0	156	16 O34598	O34598 bacillus su
14	28	80.0	179	16 Q9KBA2	Q9KBA2 bacillus ha
15	28	80.0	183	5 Q9NLC3	Q9NLC3 leishmania
16	28	80.0	241	5 O01464	O01464 caenorhabdi

17	28	80.0	262	2 P95777	P95777 streptococ
18	28	80.0	317	16 Q92M57	Q92M57 rhizobium m
19	28	80.0	353	5 Q9TVR2	Q9TVR2 caenorhabdi
20	28	80.0	483	16 O84690	O84690 chlamydia t
21	28	80.0	483	16 Q9PLP3	Q9PLP3 chlamydia m
22	28	80.0	501	5 Q19498	Q19498 caenorhabdi
23	28	80.0	554	6 Q02713	Q02713 sus scrofa
24	28	80.0	688	8 Q9BAS8	Q9BAS8 buxus henry
25	28	80.0	695	8 Q9BAS7	Q9BAS7 buxus harla
26	28	80.0	701	8 Q9MYL7	Q9MYL7 abutilon hy
27	28	80.0	1020	5 Q9XU82	Q9XU82 caenorhabdi
28	27	77.1	100	2 P74951	P74951 vibrio para
29	27	77.1	105	2 Q9R910	Q9R910 bacillus su
30	27	77.1	119	11 Q9R1V8	Q9R1V8 rattus norv
31	30	27	132	1 Q977R1	Q977R1 uncultured
32	27	77.1	145	10 Q9ARN6	Q9ARN6 oryza sativ
33	27	77.1	153	16 Q55862	Q55862 synchocyst
34	27	77.1	166	10 Q9AX62	Q9AX62 oryza sativ
35	27	77.1	180	8 Q96815	Q96815 skeletonema
36	27	77.1	207	10 Q41882	Q41882 zea mays (m
37	27	77.1	227	5 Q9BFA2	Q9BFA2 caenorhabdi
38	27	77.1	238	10 Q41883	Q41883 zea mays (m
39	27	77.1	240	10 Q41884	Q41884 zea mays (m
40	27	77.1	249	16 Q99RH2	Q99RH2 staphylococ
41	27	77.1	250	16 Q9KDY1	Q9KDY1 bacillus ha
42	27	77.1	267	10 Q82052	Q82052 sorghum bic
43	27	77.1	269	10 P93641	P93641 zea mays (m
44	27	77.1	274	17 Q97Y36	Q97Y36 sulfolobus
45	27	77.1	276	12 Q57125	Q57125 cottoncattail

ALIGNMENTS

RESULT 1
Q98IR5 ID Q98IR5 PRELIMINARY; PRT; 1346 AA.
AC Q98IR5;
DT 01-OCT-2001 (TREMBLrel. 18, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ML2286 PROTEIN.
GN ML2286.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT *Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.*
RT DNA Res. 7:331-338(2000).
RL EMBL; AP002999; BAB9451.1; -
DR InterPro: IPR003975; Shal_channel.
DR PRINTS: PR01497; SHALCHANNEL.
KW Complete proteome.
SQ SEQUENCE 1346 AA; 148727 MW; 7D1288A3FC26E879 CRC64;

Query Match 91.4%; Score 32; DB 16; Length 1346;

Best Local Similarity 85.7%; Pred. No. 73;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYQLALT 7
11:1111
DB 696 FYELALT 702

```

RESULT 2
ID 052357 PRELIMINARY; PRT; 370 AA.
AC 052357;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE PUTATIVE RAP.
OS Bacillus pumilus (Bacillus mesentericus).
OG Plasmid pPL10.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Streptococcus group; Bacillus.
OX NCBI_TaxID=1408;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 12140;
RX MEDLINE-76260025; PubMed-821919;
RA Lovett P.S., Duval E.J., Keggin K.M.;
RT "Bacillus pumilus plasmid pPL10: properties and insertion into
RL Bacillus subtilis 168 by transformation.*";
RN J. Bacteriol. 127:817-828(1976).
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 12140;
RA Lovett P.S., Rogers E.J.;
RT "Sequence of Bacillus pumilus plasmid pPL10.*";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF036712; AAB91475.1;
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR: 4.
KW Plasmid.
SQ SEQUENCE 370 AA; 44275 MW; 7302D7A53B8B3987 CRC64;

Query Match
Best Local Similarity 100.0%; Score 30; DB 2; Length 370;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYQAL 6
Db 256 FYQAL 261

RESULT 3
ID 097L93 PRELIMINARY; PRT; 744 AA.
AC 097L93;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT F.
OS NDHF.
OS Villarsia calthifolia.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Menyanthaceae; Villarsia.
OX NCBI_TaxID=13756;
RN 1)
RP SEQUENCE FROM N.A.
RX MEDLINE-20340987; PubMed-10877943;
RA Olmstead R.G., Kim K.J., Jensen R.K., Wagsstaff S.J.;
RT "The phylogeny of the asteridae sensu lato based on chloroplast ndhf
RL gene sequences.*";
RN Mol. Phylogenet. Evol. 16:96-112(2000).
CC -1- CATALYTIC ACTIVITY: NADH + PLASTOQUINONE - NAD(+) + PLASTOQUINOL.
DE EMBL: AF130180; AAF08136.1;
DR InterPro: IPR001750; Oxidored_q1.
DR InterPro: IPR002128; Oxidored_q1_C.
DR InterPro: IPR005156; Oxidored_q1_N.
DR Pfam: PF00361; oxidored_q1; 1.

```

```

DR Pfam: PF01010; oxidored_q1_C; 1.
DR Pfam: PF00662; oxidored_q1_N; 1.
KW Chloroplast; NAD; Oxidoreductase; Plastocinone.
SQ SEQUENCE 744 AA; 84557 MW; C98CDAFCCAD23724 CRC64;

Query Match
Best Local Similarity 71.4%; Score 30; DB 8; Length 744;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYQAL 7
Db 671 FYEIAL 677

RESULT 4
ID 097T70 PRELIMINARY; PRT; 857 AA.
AC 097T70;
DT 01-OCT-2001 (TREMblrel. 18, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE CELL WALL SURFACE ANCHOR FAMILY PROTEIN.
GN SP0082.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN 1)
RP SEQUENCE FROM N.A.
RC STRAIN-TIGR4;
RX MEDLINE-21357209; PubMed-11463916;
RA Tettelein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapfel E., Knouri H., Wolf A.M., Uterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RL pneumoniae.*";
RN Science 293:498-506(2001);
DR EMBL: AE007325; AAK74270.1; -.
DR TIGR: SP0082; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 857 AA; 90923 MW; 9877EFBEBCDE86F CRC64;

Query Match
Best Local Similarity 85.7%; Score 30; DB 16; Length 857;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FYQAL 6
Db 538 FYQAL 543

RESULT 5
ID 007360 PRELIMINARY; PRT; 111 AA.
AC 007360;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE NTFW.
GN NTFW.
OS Streptococcus sp. (strain PCC 8801 / RF-1) (Cyanothecae PCC 8801).
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.
OX NCBI_TaxID=41431;
RN 1)

```

RP SEQUENCE FROM N.A.
RC STRAIN-RP-1;
RX MEDLINE-99231861; PubMed-10217509;
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;
RT "Organization and expression of nitrogen-fixation genes in the aerobic
RT nitrogen-fixing unicellular cyanobacterium Synecchococcus sp. strain
RT RP-1";
RL Microbiology 145:743-753(1999).
DR EMBL; AF003700; AAC35197.1; -.
SQ SEQUENCE 111 AA; 12931 MW; C85B39697EFC310E CRC64;

Query Match 82.9%; Score 29; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLALT 7
|||||
Db 67 YOLALT 72

RESULT 6
O9KSR2 PRELIMINARY; PRT; 350 AA.
ID O9KSR2;
AC O9KSR2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN VC1194.
GN VC1194.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_Taxid=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EL TOR N16961 / SEROTYPE O1;
RX MEDLINE-20406833; PubMed-10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unanue L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bess S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
DR EMBL; AE004199; AAF94353.1; -.
DR TIGR; VC1194; -.
DR InterPro: IPR001623; DnaU_N.
DR SMART: SM00271; DnaU_1.
DR PROSITE: PSS0076; DnaU_2; 1.
KW Complete proteome.
SQ SEQUENCE 350 AA; 40627 MW; 6B6F035EAC79BBDE CRC64;

Query Match 82.9%; Score 29; DB 16; Length 350;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLALT 7
|||||
Db 50 YOLALT 55

RESULT 7
P78862 PRELIMINARY; PRT; 380 AA.
ID P78862;
AC P78862;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UNKNOWN PROTEIN (FRAGMENT).

OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PRT45;
RX MEDLINE-98162722; PubMed-9501991;
RA Yoshioaka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT "Identification of open reading frames in Schizosaccharomyces pombe
RT cDNAs";
RL DNA Res. 4:363-369(1997).
DR EMBL; D89212; BAA13873.1; -.
FT NON_TER 1
SQ SEQUENCE 380 AA; 43661 MW; 9AE19FB26E0A6303 CRC64;

Query Match 82.9%; Score 29; DB 3; Length 380;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||||
Db 203 FYOLALT 209

RESULT 8
O84844 PRELIMINARY; PRT; 658 AA.
ID O84844;
AC O84844;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 76.5 KDA PROTEIN.
GN CT837.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_Taxid=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-D/UW-3/CX;
RX MEDLINE-99000809; PubMed-9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis";
RL Science 282:754-759(1998).
DR EMBL; AE001356; AAC68434.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 658 AA; 76502 MW; E17FCA5063CCE717 CRC64;

Query Match 82.9%; Score 29; DB 16; Length 658;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||||
Db 217 FYOLALT 223

RESULT 9
O9PL82 PRELIMINARY; PRT; 658 AA.
ID O9PL82;
AC O9PL82;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN TC0225.
GN TC0225.
OS Chlamydia muridarum.

OC Bacteria: Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-MOPN / NICG;
RX MEDLINE-20150255; PubMed-10684935;
RA Read T.D., Brunham R.C., Shen C., Gili S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Knout H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson M., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOPN and Chlamydia
RT pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL: AE002289; MAF39097.1; -.
DR TIGR: TC0225; -.
DR InterPro: IPR002129; Pyridoxal dec.
DR PROSITE: PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.
DR Hypothetical protein; Complete proteome
SQ SEQUENCE 658 AA; 76487 MW; BDD228A01C6DBC8E CRC64;

Query Match 82.9%; Score 29; DB 16; Length 658;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQLALT 7
DB 217 FYQLALT 223

RESULT 10
Q9P7W6 PRELIMINARY; PRT; 684 AA.
AC Q9P7W6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE DNA MISMATCH REPAIR PROTEIN, MLHI HOMOLOG.
GN SPBC1703.04.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-972H-;
RA McDougall R.C., Rajandream M.A., Barrell B.G., Cadieu E., Lelaure V.,
RA Gallbert F.;
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL136536; CAB6448.1; -.
DR HSSP: P23367; IBKN.
DR InterPro: IPR002099; DNA_mis_repair.
DR InterPro: IPR003594; HATPase.C.
DR Pfam: PF01119; DNA_mis_repair; 1.
DR Pfam: PF02518; HATPase.C; 1.
DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; UNKNOWN_1.
DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; UNKNOWN_1.
SQ SEQUENCE 684 AA; 77253 MW; 7A59653B8E2944B CRC64;

Query Match 82.9%; Score 29; DB 3; Length 604;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQLALT 7
DB 507 FYQLALT 513

RESULT 11
Q9BKJ9 PRELIMINARY; PRT; 755 AA.
ID Q9BKJ9

AC Q9BKJ9;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE LYSOSOMAL INTEGRAL MEMBRANE PROTEIN IT.
GN LMP.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN 11
RP SEQUENCE FROM N.A.
RA Janssen K.-P., Schleicher M.;
RT Characterisation of CD36/LIMPIT-Homologues in Dictyostellium
RT discoideum.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF238324; AAK30040.1; -.
DR InterPro: IPR002159; CD36.
DR Pfam: PF01130; CD36; 1.
SQ SEQUENCE 755 AA; 82571 MW; 93BB639461C97547 CRC64;

Query Match 82.9%; Score 29; DB 5; Length 755;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YQLALT 7
DB 297 YQLALT 302

RESULT 12
Q9RYO4 PRELIMINARY; PRT; 785 AA.
AC Q9RYO4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ACULEACIN A ACYLASE.
GN DRA0255.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN-R1;
RX MEDLINE-20036896; PubMed-10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA McFarr K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanatnevan J.J., Lam P., McDonald L., Uterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL: AE001863; AAF12385.1; -.
DR TIGR: DRA0255; -.
DR InterPro: IPR002692; Penicill_amidase.
DR Pfam: PF01804; Penicill_amidase; 2.
KW Complete proteome.
SQ SEQUENCE 785 AA; 84042 MW; 031166EAB2852778 CRC64;

Query Match 82.9%; Score 29; DB 16; Length 785;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQLALT 7
DB 257 FYQLALT 263

RESULT 13
034598 PRELIMINARY: PRT: 156 AA.
AC 034598;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE YKOA.
GN YKOA OR YKNA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RL Devine K.M.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolojin A., Borcherdt S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cumming N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Ertlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shih B.S., Soldo B.,
RA Sorekin A., Taccori E., Takagi T., Takahashi H., Takemaru K.,
RA Tsuchiuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarelli A.,
RA Viati A., Wambutt R., Wedler E., Wedler H., Weltenegeger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
subtilis".
RL Nature 390:249-256(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ002571; CA05596.1; -;
DR EMBL: Z99110; CAB13174.1; -;
DR InterPro: IPR00125; dCMP_cyt_deam.
DR Pfam: PF00383; dCMP_cyt_deam; 1.
DR PROSITE: PS00903; CYT_DCMP_DEAMINASES; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 156 AA; 17156 MW; B6498345A96BC214 CRC64;

Query Match 80.0%; Score 28; DB 16; Length 156;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||||

Db 132 FYOVTLT 138
RESULT 14
O9KBA2 PRELIMINARY: PRT: 179 AA.
AC O9KBA2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE RNA POLYMERASE ECF-TYPE SIGMA FACTOR.
GN BH2026.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001514; BAB05745.1; -;
KW Complete proteome.
SQ SEQUENCE 179 AA; 21257 MW; E00A9640AC4D854A CRC64;

Query Match 80.0%; Score 28; DB 16; Length 179;
Best Local Similarity 71.4%; Pred. No. 90;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOLALT 7
|||||

Db 29 YRRLALT 35

RESULT 15
O9NLU3 PRELIMINARY: PRT: 183 AA.
AC O9NLU3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POSSIBLE PROLINE OXIDASE, MITOCHONDRIAL (FRAGMENT).
GN LM26.349.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL160493; CAB97967.1; -;
FT NON_TER 183
SQ SEQUENCE 183 AA; 20924 MW; 9C5FB37BB6F1AB2A CRC64;

Query Match 80.0%; Score 28; DB 5; Length 183;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOLAL 6
|||||

Db 57 FYOLAI 62

Search completed: July 15, 2002, 13:22:42

Job time: 1478 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:57:55 ; Search time 228.39 Seconds

(without alignments)
3.404 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35
Sequence: 1 FYQLALR 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	88.6	751	22	AAE09958	Methylomonas 16a n
2	30	85.7	262	21	AAE53254	Human colon cancer
3	30	85.7	313	22	AAE71826	Human olfactory re
4	30	85.7	624	21	AAE81588	Streptococcus pneu
5	29	82.9	9	14	AAE44285	Residues 132-140 o
6	29	82.9	9	18	AAE39629	Human p53 peptide
7	29	82.9	10	18	AAE22030	Antigen human p5
8	29	82.9	14	20	AAE25186	Mutant P53 tumour
9	29	82.9	35	20	AAE25185	Mutant P53 tumour
10	29	82.9	64	22	AAE76842	Human colon cancer
11	29	82.9	491	20	AAE37291	Protein which is s

12	29	82.9	769	22	AAE85849	Yeast MLH1 protein
13	29	82.9	769	22	AAE63951	Amino acid sequenc
14	29	82.9	3391	12	AAE13166	Proteins encoded b
15	28	80.0	247	20	AAE37489	Chlamydia trachoma
16	28	80.0	247	22	AAE52822	Escherichia coli p
17	28	80.0	480	21	AAE79308	Synechocystis sp.
18	28	80.0	1020	22	AAE01856	C. elegans RNA int
19	27	77.1	16	21	AAE84247	Consensus site for
20	27	77.1	22	19	AAE50627	Human neurofilamen
21	27	77.1	124	22	AAE75187	Human colon cancer
22	27	77.1	240	11	AAE04752	Amino acid sequenc
23	27	77.1	240	11	AAE24228	Human vesicle asso
24	27	77.1	240	22	AAE81378	Human AEP protein
25	27	77.1	240	22	AAE62393	Human type I membr
26	27	77.1	267	21	AAE42834	Human ORFX ORF2598
27	27	77.1	278	20	AAE43933	Human protein kina
28	27	77.1	278	21	AAE58121	Lung cancer associ
29	27	77.1	303	16	AAE85121	CDK4 R24C mutant.
30	27	77.1	303	20	AAE24747	Mouse CDK-4 protei
31	27	77.1	303	21	AAE96568	Cyclin-dependent k
32	27	77.1	303	21	AAE88358	Human cyclin depen
33	27	77.1	354	22	AAE12211	Human organic cati
34	27	77.1	354	22	AAE80235	Human protein SEQ
35	27	77.1	399	21	AAE30484	Arabidopsis thalia
36	27	77.1	431	21	AAE30484	Arabidopsis thalia
37	27	77.1	484	20	AAE35314	Chlamydia pneumoni
38	27	77.1	485	20	AAE28682	Human pp392.3 secr
39	27	77.1	508	22	AAE93037	Human secreted pro
40	27	77.1	508	21	AAE79307	Escherichia coli y
41	27	77.1	555	22	AAE79251	Human protein SEQ
42	27	77.1	557	21	AAE30483	Arabidopsis thalia
43	27	77.1	618	18	AAE32105	Rat sodium/iodide
44	27	77.1	618	18	AAE18571	Human cyclin D1-hu
45	27	77.1	618	19	AAE74572	Human cyclin D1/cy

ALIGNMENTS

RESULT	1
AAE09958	AAE09958 standard; Protein; 751 AA.
ID	AAE09958
AC	AAE09958;
DT	29-NOV-2001 (first entry)
DE	Methylomonas 16a nitric oxide reductase NorZ.
KW	Denitrification; enzyme; bacteria; fermentation; nitric oxide reductase;
KW	NorZ.
XX	
OS	Methylomonas sp. 16a.
XX	
PN	WO200164898-A2.
XX	
PD	07-SEP-2001.
XX	
PF	22-FEB-2001: 2001WO-US05900.
XX	
PR	29-FEB-2000: 2000US-0185621.
XX	
PA	(DUPO) DU PONT DE NEMOURS & CO E I.
XX	
PI	Odom JM, Norton KC, Schenzle AJ, Ye RW;
XX	
DR	WPI: 2001-557773/62.
XX	
DR	N-PDB: AAD17012.
XX	
PT	Novel nucleic acid fragments from Methylomonas species for use in
XX	bacterial denitrification process, encodes bacterial nitrite reductase,
XX	bacterial nitric oxide reductase or bacterial nitrate reductase

PS Claim 17: Page 60-62; 64pp; English.

XX The invention relates to a nucleic acid fragment from *Methylobomonas*

CC encoding nitrite reductase, nitric oxide reductase or nitrate reductase.

CC The nucleic acid fragment is isolated from a gene cluster containing 11

CC open reading frames (ORFs) encoding enzymes involved in denitrification.

CC The enzymes are useful in microbial denitrification reactions, for

CC identifying other denitrifying bacteria and for fermentation processes

CC in absence or presence of oxygen. Chimeric genes comprising the

CC polynucleotide of the invention can be used to alter the level of

CC denitrifying enzyme in the recombinant host. The polynucleotide is

CC useful for isolating genes encoding homologous proteins from other

CC microbial species, for expression of various pathway intermediates, for

CC modulation of pathway already existing in the host for the synthesis of

CC new products, and for conferring higher growth yields on the host or for

CC enabling alternative growth mode to be utilised. The present sequence is

CC nitric oxide reductase No12 from *Methylobomonas* 18a.

XX

SQ Sequence 751 AA;

Query Match 88.6%; Score 31; DB 22; Length 751;

Best Local Similarity 85.7%; Pred. No. 94;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQALAT 7

Db 501 fyqlglt 507

RESULT 2

AAB53254

ID AAB53254 standard; Protein: 262 AA.

XX

AC AAB53254;

XX

DT 09-MAR-2001 (first entry)

XX

DE Human colon cancer antigen protein sequence SEQ ID NO:794.

XX

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnary;

KW immunomodulatory; muscular; gastrological; gastrointestinal;

KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder.

XX

OS Homo sapiens.

XX

PN WO20005351-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05883.

XX

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

P1 Rosen CA, Ruben SM;

XX

DR WPI: 2000-587534/55.

DR N-PSDB: AAC98011.

XX

PT Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

PT disorders such as colon cancer -

XX

PS Claim 11: Page 1350-1351; 2104pp; English.

CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,

CC called human colon cancer antigens, given in AAB53254 to AAB54006. The

CC human colon cancer antigens can have cytostatic, cardioactive, muscular;

CC neuroprotective, immunomodulatory, gastrological, gastrointestinal,

CC vulnary, nephrotropic, antiinfective and antibacterial activities, and

CC can be used in gene therapy. The colon cancer antigen polynucleotides,

CC proteins and antibodies to the proteins are useful for the prevention,

CC treatment and diagnosis of colon disorders, such as colon cancer. The

CC polynucleotides may be used in diagnostics and research, such as for

CC chromosome identification, and as hybridisation probes. The proteins

CC may also be used to prevent diseases such as neural disorders, immune

CC system disorders, muscular disorders, reproductive disorders,

CC gastrointestinal disorders, wounds, renal disorders, infectious

CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and

CC AAB54007 represent sequences used in the exemplification of the present

CC invention.

XX

SQ Sequence 262 AA;

Query Match 85.7%; Score 30; DB 21; Length 262;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQAL 6

Db 19 fyqlal 24

RESULT 3

AAG71826

ID AAG71826 standard; Protein: 313 AA.

XX

AC AAG71826;

XX

DT 31-JUL-2001 (first entry)

XX

DE Human olfactory receptor polypeptide, SEQ ID NO: 1507.

XX

KW Human; olfactory receptor; OR; primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation.

XX

OS Homo sapiens.

XX

PN WO200127158-A2.

XX

PD 19-APR-2001.

XX

PF 06-OCT-2000; 2000WO-US27582.

XX

PR 08-OCT-1999; 99US-0158615.

XX

PR 24-FEB-2000; 2000US-0184809.

XX

PA (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

XX

P1 Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

XX

DR WPI: 2001-290713/30.

XX

PT New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists -

XX

PS Claim 11: Page 959-960; 1857pp; English.

CC

CC The present sequence is an olfactory receptor which is encoded by

CC one of a number of novel polynucleotides. The polynucleotides can be

CC used in screening for olfactory agonists and antagonists. The methods

CC allow for the determination of primary scents and the identification

CC of the odour receptors used to detect these primary scents. The methods

CC also enable determination of secondary scents and the identification of

CC combinations of odour receptors that are involved in detecting such

CC secondary scents. This enables the construction of a scent representation

CC (also called a scent fingerprint or scent profile), which may be used to

CC re-create and edit scents. Libraries of olfactory receptors are useful
CC for determining the interaction pattern of a composition with the
CC receptors, and can be used for determining differences in the olfactory
CC faculties of different individuals.

CC Sequence 313 AA;

Query Match 85.7%; Score 30; DB 22; Length 313;

Best Local Similarity 100.0%; Pred. No. 62;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQIAL 6

DB 197 FYQIAL 202

RESULT 4

AAV81588
ID AAV81588 standard; Protein; 624 AA.

XX AAV81588;

XX 24-MAY-2000 (first entry)

DE Streptococcus pneumoniae type 4 protein sequence #88.

KW Streptococcus pneumoniae; vaccine; screening; protein antigen;

KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;

XX Streptococcus pneumoniae.

OS Streptococcus pneumoniae.

PN WO200006737-A2.

PD 10-FEB-2000.

PF 27-JUL-1999; 99WO-GB02451.

PR 27-JUL-1998; 98GB-0016337.

PR 19-MAR-1999; 99US-0125164.

XX (MICR-) MICROBIAL TECHNICS LTD.

PI Gilbert CFG, Hansbro PM.

DR WPI; 2000-195300/17.

PT New Streptococcal protein, useful as a vaccine, for diagnosis of

PT pneumococcal diseases and for screening agents capable of antagonizing

XX or inhibiting expression of the protein

PS Claim 1; Page 81; 108pp; English.

XX AAV81501 to AAV81679 represent specifically claimed protein sequences
CC isolated from Streptococcus pneumoniae. AAA05407 to AAA05590 represent
CC specifically claimed nucleotide sequences isolated from S. pneumoniae.
CC The sequences have antibacterial and antiinflammatory properties.
CC The protein sequences, and fragments of them, are useful as immunogens
CC and/or antigens. The nucleotide sequences can be used in vaccines and in
CC diagnostic assays. The proteins and nucleotides can be useful for the
CC detection and diagnosis of S. pneumoniae. The protein sequences are also
CC useful for screening an agent capable of antagonising, inhibiting or
CC interfering with the function or expression of the proteins in which the
CC agent is useful for treatment or prophylaxis of S. pneumoniae infection
CC and meningitis. AAA05591 to AAA05614 represent primers used in the
CC exemplification of the present invention.

XX Sequence 624 AA;

Query Match 85.7%; Score 30; DB 21; Length 624;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYQIAL 6

DB 538 FYQIAL 543

RESULT 5

AAR44285
ID AAR44285 standard; peptide; 9 AA.

XX AAR44285;

XX 09-JUN-1994 (first entry)

DE Residues 132-140 of human mp53.

KW Human; p53; protein; class I; allele; diagnosis; assay; detection;

KW p53 protein-specific; T-cell; antibody; cancer; MHC; HLA-A2.1.

XX Homo sapiens.

PN WO9324525-A.

PD 09-DEC-1993.

PF 18-MAY-1993; 93WO-NL00102.

PR 26-MAY-1992; 92EP-0201510.

PA (UYLE-) RIJKSUNIV LEIDEN.

PA (SEED-) SEED CAPITAL INVESTMENTS (SCI) BV.

PI Kast WM, Melief CJM;

DR WPI; 1993-405730/50.

PT Peptide derived from p53 protein - used to treat and diagnose

PT diseases involving over-expression of p53 e.g. human cancers

XX Claim 3; Page 38; 46pp; English.

CC The sequences given in AAR44281-94 are peptides derived from human mp53

CC protein which have the ability to bind to a human Class I molecule.

CC These peptides may be used in a diagnostic test or assay to detect

CC human p53 protein-specific T-cells or antibodies. They may also be

CC used in the treatment of diseases such as human cancers showing p53

CC protein overexpression. These peptides can bind human MHC Class I

CC allele HLA-A2.1.

XX Sequence 9 AA;

OY 1 FYQIAL 7

DB 3 FYQIAL 9

RESULT 6

AAW39629
ID AAW39629 standard; peptide; 9 AA.

XX AAW39629;

XX 11-JUN-1998 (first entry)

DE Human p53 peptide #6 (pos. 132-140).

XX T cell epitope; immune response; human leukocyte antigen; HLA Class I;

KW vaccine; immunogenic; major histocompatibility complex; MHC; B cell;
 XX disease; anti-tumour; anti-viral.
 OS Homo sapiens.
 XX W09741440-A1.
 XX 06-NOV-1997.
 XX 28-APR-1997; 97WO-NL00229.
 XX 23-DEC-1996; 96EP-0203670.
 PR 26-APR-1996; 96EP-0201145.
 XX (VLE-) RIJKSUNIV LEIDEN.
 PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.
 XX Kaat WM, Melief CJM, Offringa R, Toes REM, Van Der Burg SH;
 PI WPI: 1997-549891/50.
 DR Method of selecting T cell peptide epitope(s) - by measuring the
 XX stability of HLA class I-peptide complexes on intact B cells
 PT Example 3; Page 76; 109pp; English.
 XX Peptides AAW39430-W39734 are used in a novel method for the selection of
 CC immunogenic T-cell peptide epitopes present in polypeptide antigens. The
 CC method involves the identification of peptide sequences capable of
 CC binding to an HLA (human leukocyte antigen) class I molecule and
 CC measuring the binding of this epitope peptide to the HLA class I
 CC peptide. The stability of binding of the peptide and MHC (major
 CC histocompatibility complex) class I molecule is measured on intact human
 CC B cells carrying the MHC molecule at their cell surfaces. The method can
 CC be used to select peptide epitopes for generating vaccines against a
 CC disease associated with the polypeptide, e.g. cancers or AIDS. The
 CC peptide epitopes are especially T-cell peptide epitopes with strong
 CC anti-tumour and anti-viral immune responses. Peptide AAW39629 is derived
 CC from human p53 and is capable of binding to the human MHC Class I allele
 CC HLA-A2.1.
 XX SQ Sequence 9 AA;
 Query Match 82.9%; Score 29; DB 18; Length 9;
 Best Local Similarity 85.7%; Pred. No. 6.4e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FYQALAT 7
 Db 3 fytqlakt 9
 RESULT 7
 AAW22030
 ID AAW22030 standard; peptide: 10 AA.
 XX AAW22030;
 AC AAW22030;
 XX 20-FEB-1998 (first entry)
 DT Antigenic human P53 peptide V10.
 DE Antigenic human P53 peptide V10.
 XX Antigenic peptide; human papillomavirus; human P53; V10; MAGE gene;
 KW human immunodeficiency virus; cancer antigen; tyrosinase; signal protein;
 KW anthrax lethal factor; LF; toxin; cationic fusion peptide; translocation;
 KW gene therapy; polycationic affinity handle; therapeutic protein; LFN.
 XX Homo sapiens.
 OS Homo sapiens.
 XX W09723236-A1.
 PN W09723236-A1.
 XX 03-JUL-1997.

XX 13-DEC-1996; 96WO-US20463.
 PF 07-JUN-1996; 96US-0019275.
 XX 13-DEC-1995; 95US-0008518.
 PR (HARD) HARVARD COLLEGE.
 XX Ballard JD, Blanke SR, Collier RJ, Lysak EL, Milne JC;
 PI Starnbach MN;
 DR WPI: 1997-350782/32.
 XX Introducing therapeutic proteins, especially antigens, into cells -
 PT using toxin molecules and/or polycationic handles for delivery
 XX Claim 15; Page 36; 67pp; English.
 XX This is the antigenic human P53 peptide V10. This antigenic compound
 CC can be introduced into the cytoplasm of a cell by a new method where
 CC the cell is contacted with a fusion molecule comprising a delivery
 CC molecule. The delivery molecule can either be a polycationic affinity
 CC handle, LFN (the protective antigen binding domain of anthrax lethal
 CC factor) or a toxin delivery molecule related to LFN. The antigenic
 CC compound is linked to either of the delivery molecules by a covalent
 CC bond. The B moiety of a toxin enhances delivery of the antigenic compound
 CC into a cell. The anthrax toxin system of the invention eliminates the
 CC need to generate fusion proteins with a toxin B moiety, which alleviates
 CC problems associated with incorrect folding of lengthy fusion proteins.
 CC Small cationic fusion peptides substituted for LFN may reduce the
 CC possibility of steric interference with the biological activity of the
 CC translocated protein. The method is used for the introduction of
 CC antigens, e.g. MHC class I antigens or any other therapeutic protein,
 CC e.g. toxin molecules, apoptosis-inducing molecules or signalling
 CC proteins into the cells.
 XX SQ Sequence 10 AA;
 Query Match 82.9%; Score 29; DB 18; Length 10;
 Best Local Similarity 85.7%; Pred. No. 2.5;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 FYQALAT 7
 Db 1 fytqlakt 7
 RESULT 8
 AAY25186
 ID AAY25186 standard; peptide: 14 AA.
 XX AAY25186;
 AC AAY25186;
 XX 03-SEP-1999 (first entry)
 DT Mutant P53 tumour suppressor protein peptide #2.
 DE Heat shock protein; HSP; complex; denatured protein matrix; antigen;
 KW vaccine; allergic disease; treatment; susceptibility; Th2; skin rash;
 KW allergic reaction; asthma; P53; tumour suppressor protein; mutant.
 XX Synthetic.
 OS Synthetic.
 XX W09929182-A1.
 PN W09929182-A1.
 XX 17-JUN-1999.
 PD 04-DEC-1998; 98WO-US25734.
 XX 05-DEC-1997; 97US-0986234.
 PR 05-DEC-1997; 97US-0985548.
 XX

PA (UYNE-) UNIV NEW MEXICO STATE.
XX
PI Moseley PL, Wallen ES;
XX
DR WPI, 1999-394912/33.
XX
PT Synthesizing heat shock protein complexes using a denatured protein
matrix
XX
PS Example 1; Fig 1A; 33pp; English.
XX
CC This invention describes a novel method for synthesizing heat shock
CC protein (HSP) complexes comprising adding a heat shock protein to a
CC denatured protein matrix for binding, and adding a complexing solution
CC comprising a peptide to elute a heat shock protein-peptide complex. A
CC HSP-antigen complex is useful as a vaccine for treating an allergic
CC disease (in a mammal, preferably a human) to reduce susceptibility of
CC the Th2 response, the complex comprising a HSP-antigenic peptide complex.
CC The complex is administered to prevent a mammal from having an allergic
CC reaction to an allergic disease, or administered to a mammal having an
CC allergic disease, to reduce the allergic reactions. Allergic diseases
CC include asthma and skin rashes. Prior art methods or preventing/treating
CC allergic diseases include antihistamines which treat only the symptoms,
CC corticosteroids which have severe side effects and desensitization
CC therapy which has limited uses. The new method also allows more
CC flexibility of use of peptide-based vaccines, as prior art HSP-based
CC vaccines require isolation from a portion of the tumour itself. This
CC sequence represents a peptide fragment from the p53 tumour suppressor
CC protein which is used in the method of the invention.
XX
SQ Sequence 14 AA;

Query Match 82.9%; Score 29; DB 20; Length 14;
Best Local Similarity 85.7%; Pred. No. 3.6;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
| | | | |
Db 3 fyqlakt 9

RESULT 9
AAV25185
ID AAV25185 standard; peptide; 35 AA.
AC AAV25185;
XX
DT 03-SEP-1999 (first entry)
XX
DE Mutant p53 tumour suppressor protein peptide #1.
XX
KW Heat shock protein; HSP; complex; denatured protein matrix; antigen;
KW vaccine; allergic disease; treatment; susceptibility; Th2; skin rash;
KW allergic reaction; asthma; p53; tumour suppressor protein; mutant.
XX
OS Synthetic.
XX
PN WO9929182-A1.
XX
PD 17-JUN-1999.
XX
PF 04-DEC-1998; 98WO-US25734.
XX
PR 05-DEC-1997; 97US-0986234.
PR 05-DEC-1997; 97US-0985548.
XX
PA (UYNE-) UNIV NEW MEXICO STATE.
XX
PI Moseley PL, Wallen ES;
XX
DR WPI, 1999-394912/33.
XX

PT Synthesizing heat shock protein complexes using a denatured protein
matrix
XX
PS Example 1; Fig 1A; 33pp; English.
XX
CC This invention describes a novel method for synthesizing heat shock
CC protein (HSP) complexes comprising adding a heat shock protein to a
CC denatured protein matrix for binding, and adding a complexing solution
CC comprising a peptide to elute a heat shock protein-peptide complex. A
CC HSP-antigen complex is useful as a vaccine for treating an allergic
CC disease (in a mammal, preferably a human) to reduce susceptibility of
CC the Th2 response, the complex comprising a HSP-antigenic peptide complex.
CC The complex is administered to prevent a mammal from having an allergic
CC reaction to an allergic disease, or administered to a mammal having an
CC allergic disease, to reduce the allergic reactions. Allergic diseases
CC include asthma and skin rashes. Prior art methods or preventing/treating
CC allergic diseases include antihistamines which treat only the symptoms,
CC corticosteroids which have severe side effects and desensitization
CC therapy which has limited uses. The new method also allows more
CC flexibility of use of peptide-based vaccines, as prior art HSP-based
CC vaccines require isolation from a portion of the tumour itself. This
CC sequence represents a peptide fragment from the p53 tumour suppressor
CC protein which is used in the method of the invention.
XX
SQ Sequence 35 AA;

Query Match 82.9%; Score 29; DB 20; Length 35;
Best Local Similarity 85.7%; Pred. No. 9.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
| | | | |
Db 10 fyqlakt 16

RESULT 10
AAG76842
ID AAG76842 standard; Protein; 64 AA.
XX
AC AAG76842;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:7606.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 21.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI, 2001-235357/24.
DR N-PSDB; AAH36247.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 9011-9013; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patient's own production of P.
CC Additionally, N may be used to produce the colon cancer-associated ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
CC
XX
SQ Sequence 64 AA;

Query Match 82.9%; Score 29; DB 22; Length 64;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 YQALAT 7
|||
|||
Db 38 yqlalt 43

RESULT 11

AAV37291
ID AAV37291 standard; Protein: 491 AA.

XX AAV37291;

DT 07-OCT-1999 (first entry)

DE Protein which is specific to Chlamydia trachomatis.

XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; peritropatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB01939.
XX
PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97ER-0015041.
PR 17-DEC-1997; 97FR-0016034.
XX
PA (GEST) GENSET.
XX
PI Griffiths R;
XX
DR WPI: 1999-371125/31.
XX
PT Genome sequence of Chlamydia trachomatis
XX
PS Disclosure; Page #VALUE1; 1755pp; English.
XX
XX AAV36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AA01425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye

CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis;
CC peritropatitis, bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.
XX
SQ Sequence 491 AA;

Query Match 82.9%; Score 29; DB 20; Length 491;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQALAT 7
|||
|||
Db 49 fyqkalt 55

RESULT 12

AAB5849
ID AAB5849 standard; Protein: 769 AA.

XX AAB5849;

DT 29-OCT-2001 (first entry)

DE Yeast MLH1 protein.

XX Hypermutable bacteria; mismatch repair gene; MMR gene; Muth; Muts; Mutt;
KW Muth; PMS2; MLH1; MLH3; PMSR; biocatalysis; bioremediation; biochemical;
KW drug discovery; detoxification; toxin; biotransformation.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200159092-A2.
XX
PD 16-AUG-2001.
XX
PF 12-FEB-2001; 2001WO-US04339.
XX
PR 11-FEB-2000; 2000US-0181929.
XX
PA (UYUO) UNIV JOHNS HOPKINS.
XX
PI Nicolaiides NC, Sass PM, Grasso L, Vogelstein B, Kinzler KW;
XX
XX WPI: 2001-514664/56.
XX
DR N-PSDB; AAH76362.
XX

PT Making hypermutable bacteria for biocatalysis, bioremediation and drug
PT discovery, involves introducing polynucleotide comprising dominant
PT negative allele of mismatch repair gene under regulatory sequence
PT control -
XX

XX Example 1; Page 37-38; 68pp; English.

XX The invention provides a method for generating a hypermutable bacteria.
CC The method involves introducing a polynucleotide having a dominant
CC negative allele of a mismatch repair (MMR) gene under the control of an
CC inducible transcription regulatory sequence, into a bacterium. The cell
CC becomes inducibly hypermutable. The method is useful to create desirable
CC output traits for commercial applications, using dominant negative
CC alleles of mismatch repair proteins. The mismatch repair gene is a Muth,
CC Muts, Mutt or Muth homologue and can be selected from PMS2, MLH1, MLH3,
CC PMSR or PMSR homologue. The hypermutable bacteria is useful for the
CC production, biocatalysis, bioremediation and drug discovery. It is also
CC useful in manufacturing industry for the generation of new biochemicals
CC useful for detoxifying noxious chemicals from by-products of
CC manufacturing processes or those used as catalysts, for remediation of
CC toxins present in the environment including polychlorobenzene, heavy
CC metals and other environmental hazards for which there is a need to
CC remove them from the environment. The hypermutable bacteria is also

CC useful for screening novel mutations in a gene or a set of genes that
CC produce variant siblings that exhibit a new output trait not found in
CC wild type cells. The bacteria are also useful for producing increased
CC quantity or quality of protein or non-protein therapeutic molecule e.g.
CC penicillin G, Erythromycin and Clavulanic acid, by biotransformation.
CC Dominant negative alleles of the MMR gene are useful for producing higher
CC quantities of recombinant polypeptides. The present sequence represents
CC a yeast MLH1 protein.
XX
SQ Sequence 769 AA;

Query Match 82.9%; Score 29; DB 22; Length 769;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
|||: ||
Db 570 fyqigt 576

RESULT 13

AA63951
ID AAG63951 standard; Protein: 769 AA.

XX
AC AAG63951;

XX 29-OCT-2001 (first entry)

XX Amino acid sequence of yeast mismatch repair protein MLH1.

XX MLH1: mismatch repair gene; MMR gene; hypermutable yeast.

OS Unidentified.

XX WO200162945-A1.

XX 30-AUG-2001.

XX 21-FEB-2001; 2001WO-US05447.

XX 23-FEB-2000; 2000US-0184336.

XX (UYJO) UNIV JOHNS HOPKINS.

XX (NICO/) NICOLAIDES N C.

XX (SASS/) SASS P M.

XX (GRAS/) GRASSO L.

XX (VOGE/) VOGELSTEIN B.

XX (KINZ/) KINZLER K W.

XX NICOLAIDES NC, SASS PM, GRASSO L, Vogelstein B, Kinzler KW;

XX WPI, 2001-522820/57.

XX N-PSDB; AAH75039.

XX Disclosure; Page 36; 60pp; English.

XX The present sequence represents yeast MLH1. MLH1 is a mismatch repair
CC (MMR) gene. The specification describes a method for making a
CC hypermutable yeast, comprising introducing a polynucleotide containing
CC a dominant negative allele of a mismatch repair (MMR) gene, into a
CC yeast, whereby the cell becomes hypermutable. The method is useful
CC to create desirable output traits for commercial applications, using
CC dominant negative alleles of mismatch repair proteins. The hypermutable
CC yeast is useful for production, biocatalysis, bioremediation and drug
CC discovery. It is also useful in genetic screens for the direct selection
CC of variant subclones that exhibit new output traits. The hypermutable
CC yeast is also useful in the manufacturing industry for the generation
CC of new biochemicals, for detoxifying noxious chemicals from by-products

CC of manufacturing processes or those used as catalysts, for remediation
CC of toxins present in the environment including polychlorobenzenes, heavy
CC metals and other environmental hazards for which there is a need to
CC remove them from the environment. The yeast is also useful for producing
CC increased quantity or quality of protein or non-protein therapeutic
CC molecule e.g., Penicillin G, Erythromycin and Clavulanic acid, by
CC biotransformation.
XX
SQ Sequence 769 AA;

Query Match 82.9%; Score 29; DB 22; Length 769;
Best Local Similarity 71.4%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FYQLALT 7
|||: ||
Db 570 fyqigt 576

RESULT 14

AAR13166
ID AAR13166 standard; Protein: 3391 AA.

XX
AC AAR13166;

XX 21-NOV-1991 (first entry)

XX Proteins encoded by entire Dengue 2 virus genome.

XX dengue virus; detection; consensus sequence; Flavivirus; PCR.

XX Dengue virus.

XX Key Location/Qualifiers

XX Peptide 116..205

XX Protein 206..280

XX Protein 281..775

XX Protein 776..1127

XX Protein 1128..1345

XX Protein 1346..1474

XX Protein 1475..2093

XX Protein 2094..2243

XX Protein 2244..2492

XX Protein 2493..3391

XX Protein 183

XX Modified-site 347

XX Modified-site 433

XX Modified-site 905

XX Modified-site 982

XX Modified-site 1134

XX Modified-site 1174

XX Modified-site 1329

XX Modified-site 1369

XX Modified-site 1369

OS	Chlamydia trachomatis.
XX	
PN	W09928475-A2.
XX	
PD	10-JUN-1999.
XX	
FE	27-NOV-1998; 98WO-1B01939.
XX	
PR	04-NOV-1998; 98US-0107077.
XX	
PR	28-NOV-1997; 97FR-0015041.
XX	
PR	17-DEC-1997; 97ER-0016034.
XX	
PA	(GEST) GENSET.
XX	
PI	Griffats R;
XX	
DR	WPI; 1999-371125/31.
XX	
FT	Genome sequence of Chlamydia trachomatis
XX	
PS	Disclosure; Page 1170; 1755pp; English.
XX	
XX	AAV36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC	of Chlamydia trachomatis (see AA01425). The polypeptides can be used as
CC	vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC	can also be used to control growth of the microorganism. Chlamydia
CC	trachomatis is responsible for a large number of diseases, e.g. eye
CC	diseases such as conventional trachoma, nongonococcal trachoma,
CC	paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC	nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC	peritonitis, Bartholinitis; pneumonia in breast feeding infants;
CC	and venereal lymphogranulomatosis. The polypeptides of the invention
CC	may be of use in treating these diseases.
XX	
XX	Sequence 247 AA;

Query Match	80.0%;	Score 28;	DB 20;	Length 247;
Best Local Similarity	71.4%;	Pred. No. 1.4e+02;		
Matches	5;	Conservative	1;	Mismatches 1; Indels 0; Gaps 0

Qy	1	FXQLALT	7
	11	:	1111
Db	167	fysialt	173

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:59:17 ; Search time 75.67 Seconds

(without alignments)
2.260 Million cell updates/sec

Title: US-09-712-819A-3

Perfect score: 35

Sequence: 1 FYOLAL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_A1:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTOUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	30	85.7	502	4	US-09-413-814-106
2	29	82.9	9	1	US-08-338-634-30
3	29	82.9	10	2	US-08-528-129A-5
4	29	82.9	12	2	US-08-528-129A-3
5	29	82.9	14	2	US-08-986-234-10
6	29	82.9	14	2	US-08-528-129A-4
7	29	82.9	21	2	US-08-528-129A-1
8	29	82.9	35	2	US-08-986-234-9
9	29	82.9	770	4	US-08-209-521-13
10	29	82.9	770	4	US-08-961-810-123
11	29	82.9	770	4	US-08-352-902D-123
12	27	77.1	278	2	US-07-857-224B-32
13	27	77.1	303	2	US-08-306-511A-9
14	27	77.1	303	2	US-08-893-274-9
15	27	77.1	303	3	US-08-581-918A-9
16	27	77.1	303	4	US-08-346-147B-9
17	27	77.1	303	4	US-08-822-936-9
18	27	77.1	303	5	PCT-US95-04636-9
19	27	77.1	553	3	US-08-501-572-3
20	27	77.1	553	3	US-09-040-444-3
21	27	77.1	618	2	US-08-770-761A-3
22	27	77.1	647	2	US-08-770-761A-8
23	27	77.1	660	2	US-08-770-761A-2
24	27	77.1	662	2	US-08-770-761A-5
25	27	77.1	705	2	US-08-770-761A-7
26	26	74.3	9	2	US-08-986-234-11
27	26	74.3	363	6	5223606-6

28	26	74.3	765	4	US-08-444-818-70	Sequence 70, Appl
29	26	74.3	933	4	US-08-764-870-14	Sequence 14, Appl
30	26	74.3	933	4	US-08-980-115-14	Sequence 14, Appl
31	25	71.4	65	1	US-08-055-917-14	Sequence 14, Appl
32	25	71.4	65	1	US-08-055-917-16	Sequence 16, Appl
33	25	71.4	65	1	US-08-055-917-18	Sequence 18, Appl
34	25	71.4	65	1	US-08-095-068-14	Sequence 14, Appl
35	25	71.4	65	1	US-08-095-068-16	Sequence 16, Appl
36	25	71.4	65	1	US-08-095-068-18	Sequence 18, Appl
37	25	71.4	65	1	US-08-140-721A-14	Sequence 14, Appl
38	25	71.4	65	1	US-08-140-721A-16	Sequence 16, Appl
39	25	71.4	65	1	US-08-140-721A-18	Sequence 18, Appl
40	25	71.4	65	1	US-08-619-790C-14	Sequence 14, Appl
41	25	71.4	65	1	US-08-619-790C-16	Sequence 16, Appl
42	25	71.4	65	1	US-08-619-790C-18	Sequence 18, Appl
43	25	71.4	65	2	US-07-785-565A-14	Sequence 14, Appl
44	25	71.4	65	2	US-07-785-565A-16	Sequence 16, Appl
45	25	71.4	65	2	US-07-785-565A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-413-814-106
; Sequence 106, Application US/09413814
; Patent No. 6225064
GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard
; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or
TITLE OF INVENTION: heteropolypeptide compounds
FILE REFERENCE: PCT/US 99/23535
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: US/09/413,814
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 106
LENGTH: 502
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-106

Query Match
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FYOLAL 6
DB 20 FYOLAL 25

RESULT 2
US-08-338-634-30
; Sequence 30, Application US/08338634
; Patent No. 5679641
GENERAL INFORMATION:
; APPLICANT:
TITLE OF INVENTION: Peptides of human p53 protein for use
TITLE OF INVENTION: in human T cell response inducing compositions, and
TITLE OF INVENTION: human p53 protein-specific cytotoxic T-lymphocytes.

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338, 634
FILING DATE: 06-February-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL93/00102
FILING DATE: 18-May-1993
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29, 281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3530
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-338-634-30

Query Match 82.9%; Score 29; DB 1; Length 9;
Best Local Similarity 85.7%; Pred. No. 1.7e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0;

Oy 1 FVQLALT 7
 |||||
Db 3 FVQLAKT 9

RESULT 3
US-08-528-129A-5
Sequence 5, Application US/08528129A
Patent No. 5997869
GENERAL INFORMATION:
APPLICANT: Goletz, Theresa J.
APPLICANT: Berzofsky, Jay A.
APPLICANT: Helman, Lee J.
TITLE OF INVENTION: PEPTIDES CONTAINING A FUSION JOINT OF A
TITLE OF INVENTION: CHIMERIC PROTEIN ENCODED BY DNA SPANNING A TUMOR-ASSOCIATED CH
TITLE OF INVENTION: TRANSLOCATION AND THEIR USE AS IMMUNOGENS (as amended)
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528, 129A
FILING DATE: 14-SEP-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/424, 573
FILING DATE: 17-APR-1995
APPLICATION NUMBER: 08/031,494
FILING DATE: 15-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D, John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 08830/012001
REFERENCE/DOCKET NUMBER: DHHS Ref. No. 5997869 E-220-95/0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-528-129A-5

Query Match 82.9%; Score 29; DB 2; Length 10;
Best Local Similarity 85.7%; Pred. No. 0.39;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 FVQLALT 7
 |||||
Db 1 FVQLAKT 7

RESULT 4
US-08-528-129A-3
Sequence 3, Application US/08528129A
Patent No. 5997869
GENERAL INFORMATION:
APPLICANT: Goletz, Theresa J.
APPLICANT: Berzofsky, Jay A.
APPLICANT: Helman, Lee J.
TITLE OF INVENTION: PEPTIDES CONTAINING A FUSION JOINT OF A
TITLE OF INVENTION: CHIMERIC PROTEIN ENCODED BY DNA SPANNING A TUMOR-ASSOCIATED
TITLE OF INVENTION: TRANSLOCATION AND THEIR USE AS IMMUNOGENS (as amended)
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528, 129A
FILING DATE: 14-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/424, 573
FILING DATE: 17-APR-1995
APPLICATION NUMBER: 08/031,494
FILING DATE: 15-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D, John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 08830/012001
REFERENCE/DOCKET NUMBER: DHHS Ref. No. 5997869 E-220-95/0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-528-129A-3

Query Match 82.9%; Score 29; DB 2; Length 12;
Best Local Similarity 85.7%; Pred. No. 0.48;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
DB 6 FYQALAKT 12

RESULT 5
US-08-986-234-10
Sequence 10, Application US/08986234
Patent No. 5981706

GENERAL INFORMATION:
APPLICANT: Wallen, et al.
TITLE OF INVENTION: Methods for synthesizing Heat Shock Protein Complexes
FILE REFERENCE: UNNE-0008-1
CURRENT APPLICATION NUMBER: US/08/986,234
CURRENT FILING DATE: 1997-12-05
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 14
TYPE: PRT
ORGANISM: human
US-08-986-234-10

Query Match 82.9%; Score 29; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.58;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
DB 3 FYQALAKT 9

RESULT 6
US-08-528-129A-4
Sequence 4, Application US/08528129A
Patent No. 5997869

GENERAL INFORMATION:
APPLICANT: Goletz, Theresa J.
APPLICANT: Berzofsky, Jay A.
TITLE OF INVENTION: PEPTIDES CONTAINING A FUSION JOINT OF A
TITLE OF INVENTION: CHIMERIC PROTEIN ENCODED BY DNA SPANNING A TUMOR-ASSOCIATED CH
TITLE OF INVENTION: TRANSLOCATION AND THEIR USE AS IMMUNOGENS (as amended)
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,129A
FILING DATE: 14-SEP-1995

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/424,573
FILING DATE: 17-APR-1995
APPLICATION NUMBER: 08/031,494
FILING DATE: 15-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D, John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 08830/012001
REFERENCE/DOCKET NUMBER: DHHS Ref. No. 5997869 E-220-95/0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-528-129A-4

Query Match 82.9%; Score 29; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.58;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FYQALAT 7
DB 3 FYQALAKT 9

RESULT 7
US-08-528-129A-1
Sequence 1, Application US/08528129A
Patent No. 5997869

GENERAL INFORMATION:
APPLICANT: Goletz, Theresa J.
APPLICANT: Berzofsky, Jay A.
TITLE OF INVENTION: PEPTIDES CONTAINING A FUSION JOINT OF A
TITLE OF INVENTION: CHIMERIC PROTEIN ENCODED BY DNA SPANNING A TUMOR-ASSOCIATED
TITLE OF INVENTION: TRANSLOCATION AND THEIR USE AS IMMUNOGENS (as amended)
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,129A
FILING DATE: 14-SEP-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/424,573
FILING DATE: 17-APR-1995
APPLICATION NUMBER: 08/031,494
FILING DATE: 15-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr., Ph.D, John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 08830/012001
REFERENCE/DOCKET NUMBER: DHHS Ref. No. 5997869 E-220-95/0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099

```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-528-129A-1

Query Match      82.9%; Score 29; DB 2; Length 21;
Best Local Similarity 85.7%; Pred. No. 0.92;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQLALT 7
   |||||
DB 10 FVQLAKT 16

RESULT 8
US-08-986-234-9
; Sequence 9, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Mallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNNE-0008-1
; CURRENT APPLICATION NUMBER: US/08/986,234
; CURRENT FILING DATE: 1997-12-05
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 35
; TYPE: PRT
; ORGANISM: human
US-08-986-234-9

Query Match      82.9%; Score 29; DB 2; Length 35;
Best Local Similarity 85.7%; Pred. No. 1.7;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQLALT 7
   |||||
DB 10 FVQLAKT 16

RESULT 9
US-08-209-521-13
; Sequence 13, Application US/08209521
; Patent No. 5922855
; GENERAL INFORMATION:
; APPLICANT: Liskay, Robert M.
; APPLICANT: Bronner, C. Eric
; APPLICANT: Baker, Sean M.
; APPLICANT: Bollag, Roni J.
; APPLICANT: Kolodner, Richard D.
; TITLE OF INVENTION: MAMMALIAN DNA MISMATCH REPAIR GENES
; TITLE OF INVENTION: hMLH1 AND hPMS1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heuser
; STREET: 520 S.W. Yamhill, Suite 200
; CITY: Portland
; STATE: Oregon
; COUNTRY: US
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```

; APPLICATION NUMBER: US/08/209,521
; FILING DATE: 08-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Rysselberghe, Pierre C.
; REGISTRATION NUMBER: 33,557
; REFERENCE/DOCKET NUMBER: OHSU 306A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 224-6655
; TELEFAX: (503) 295-6679
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-209-521-13

Query Match      82.9%; Score 29; DB 2; Length 770;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQLALT 7
   |||||
DB 571 FVQLIGT 577

RESULT 10
US-08-961-810-123
; Sequence 123, Application US/08961810
; Patent No. 6165713
; GENERAL INFORMATION:
; APPLICANT: Liskay, Robert M.
; APPLICANT: Bronner, C. Eric
; APPLICANT: Baker, Sean M.
; APPLICANT: Bollag, Roni J.
; APPLICANT: Kolodner, Richard D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA
; TITLE OF INVENTION: MISMATCH REPAIR GENES
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kolisch, Hartwell, Dickinson, McCormack &
; ADDRESSEE: Heuser
; STREET: 520 S.W. Yamhill Street, Suite 200
; CITY: Portland
; STATE: Oregon
; COUNTRY: U.S.A.
; ZIP: 97204
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,810
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Rysselberghe, Pierre C.
; REGISTRATION NUMBER: 33,557
; REFERENCE/DOCKET NUMBER: OHSU 306B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 224-6655
; TELEFAX: (503) 295-6679
; TELEX: 360619
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 770 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

MOLECULE TYPE: protein
US-08-961-810-123

Query Match 82.9%; Score 29; DB 4; Length 770;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQALVT 7
111:11
DB 571 FVQIGLT 577

RESULT 11

US-08-352-902D-123
Sequence 123, Application US/08352902D
Patent No. 6191268
GENERAL INFORMATION:
APPLICANT: Liskey, Robert M.
Bromner, C. Eric
Baker, Sean M.
Bollag, Roni J.
Kolodner, Richard D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATING TO DNA
MISMATCH REPAIR GENES
NUMBER OF SEQUENCES: 149
CORRESPONDENCE ADDRESS:
ADDRESSER: Kolisch, Hartwell, Dickinson, McCormack &
Heuser
STREET: 520 S.W. Yamhill Street, Suite 200
CITY: Portland
STATE: Oregon
COUNTRY: U.S.A.
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08352, 902D
APPLICATION NUMBER: US/08352, 902D
FILING DATE: 09-Dec-1994
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Van Ryselbergh, Pierre C.
REGISTRATION NUMBER: 33,557
REFERENCE/DOCKET NUMBER: OHSU 306B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 224-6655
TELEFAX: (503) 295-6679
TELEX: 360619
INFORMATION FOR SEQ ID NO: 123:
SEQUENCE CHARACTERISTICS:
LENGTH: 770 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 123:
US-08-352-902D-123

Query Match 82.9%; Score 29; DB 4; Length 770;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQALVT 7
111:11
DB 571 FVQIGLT 577

RESULT 12
US-07-857-224B-32

Sequence 32, Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857, 224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 278
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: human
FEATURE: Protein kinase; Table 8 Column 36
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988

US-07-857-224B-32

Query Match 77.1%; Score 27; DB 2; Length 278;
Best Local Similarity 83.3%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YQALVT 7
11:111
DB 164 YQALVT 169

RESULT 13
US-08-306-511A-9
Sequence 9, Application US/08306511A
Patent No. 5962316
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Demetrick, Douglas J.
APPLICANT: Seriano, Manuel
APPLICANT: Hannon, Gregory J.
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD

STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,511A
FILING DATE: 14-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSI-001CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
US-08-306-511A-9

Query Match 77.1%; Score 27; DB 2; Length 303;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLALT 7
11:111
DB 167 YOMALT 172

RESULT 14
US-08-893-274-9
Sequence 9, Application US/08893274
Patent No. 5968821
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Demetrick, Douglas J.
APPLICANT: Serrano, Manuel
APPLICANT: Hannon, Gregory J.
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses Related Thereeto
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,274
FILING DATE: 15-JULY-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,511
FILING DATE: 14-SEPTEMBER-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/248,812
FILING DATE: 25-MAY-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,371
FILING DATE: 14-APRIL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/154,915
FILING DATE: 18-NOVEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,997
FILING DATE: 17-DECEMBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/963,308
FILING DATE: 16-OCTOBER-1992
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MTV-071.09
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 303 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: N-terminal
US-08-893-274-9

Query Match 77.1%; Score 27; DB 2; Length 303;
Best Local Similarity 83.3%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 YOLALT 7
11:111
DB 167 YOMALT 172

RESULT 15
US-08-581-918A-9
Sequence 9, Application US/08581918A
Patent No. 6043030
GENERAL INFORMATION:
APPLICANT: Beach, David H.
APPLICANT: Demetrick, Douglas J.
APPLICANT: Serrano, Manuel
APPLICANT: Hannon, Gregory J.
TITLE OF INVENTION: Cell-Cycle Regulatory Proteins, and Uses Related Thereeto
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELLIOT
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPad
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/581,918A
FILING DATE: 02-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/497,214
FILING DATE: 30-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/346,147
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,511
 FILING DATE: 14-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/248,812
 FILING DATE: 25-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/227,371
 FILING DATE: 14-APR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/154,915
 FILING DATE: 18-NOV-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/991,997
 FILING DATE: 17-DEC-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MIV-071.06
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1299
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 303 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: N-terminal
 US-08-581-918A-9

Query Match 77.1%; Score 27; DB 3; Length 303;
 Best Local Similarity 83.3%; Pred No. 60;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 YQLALT 7
 11:111
 Db 167 YQMALF 172

Search completed: July 15, 2002, 12:59:17
 Job time: 388 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:01:01 ; Search time 95.45 Seconds
(without alignments)
8.054 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46

Sequence: 1 PKLIYWA 8

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	46	100.0	81	2	PH1048
2	46	100.0	92	2	S37533
3	46	100.0	92	2	S37530
4	46	100.0	92	2	S37529
5	46	100.0	92	2	S37535
6	46	100.0	92	2	S37534
7	46	100.0	92	2	S37532
8	46	100.0	96	2	G38601
9	46	100.0	98	2	S19974
10	46	100.0	101	2	S26337
11	46	100.0	101	2	PH1046
12	46	100.0	101	2	PH0869
13	46	100.0	103	2	PH1050
14	46	100.0	103	2	PH1051
15	46	100.0	103	2	PH1052
16	46	100.0	103	2	PH1054
17	46	100.0	103	2	PH1047
18	46	100.0	104	2	PH1101
19	46	100.0	104	2	PH1102
20	46	100.0	104	2	PH1103
21	46	100.0	104	2	PH1104
22	46	100.0	104	2	F38601
23	46	100.0	107	2	S36268
24	46	100.0	111	2	S03304
25	46	100.0	111	2	G30502
26	46	100.0	112	2	PI0265
27	46	100.0	112	2	S09970
28	46	100.0	112	2	S41393
29	46	100.0	112	2	S43103

30	46	100.0	113	2	PT0407	Ig light chain V r
31	46	100.0	113	2	PT0408	Ig light chain V r
32	46	100.0	113	2	S30523	Ig kappa chain V r
33	46	100.0	113	2	S30520	Ig kappa chain V r
34	46	100.0	113	2	PL0263	Ig kappa chain V r
35	46	100.0	113	2	A49260	antitumor monoclon
36	46	100.0	113	2	S34003	Ig kappa chain V r
37	46	100.0	113	2	S34002	Ig kappa chain V r
38	46	100.0	114	1	K4H0LN	Ig kappa chain V-I
39	46	100.0	114	2	S44119	Ig kappa chain V-J
40	46	100.0	117	2	S42466	Ig kappa chain V r
41	46	100.0	118	2	PT0356	Ig kappa chain V r
42	46	100.0	120	2	G33932	Ig kappa chain pre
43	46	100.0	120	2	S51147	antibody light cha
44	46	100.0	121	1	K4HU	Ig kappa chain pre
45	46	100.0	131	2	PL0207	anti-idiotypic ant

ALIGNMENTS

RESULT 1
PH1048
Ig light chain V region (clone 165.49) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Jun-1996
C:Accession: PH1048
R.Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1048
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-81 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PKLIYWA 8
DB 29 PKLIYWA 36
RESULT 2
S37533
Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37533
R.Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human TGM(+)(9D(+)) cells, the major B cell subset in the peripheral bl
A:Reference number: S37501
A:Accession: S37533
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26627; NID:9405708; PTDN:CAA81380.1; PTD:9405709
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PKLIYWA 8

Db 28 PKLLIYWA 35

RESULT 3

Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37530
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IgD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37530
A:Molecule type: mRNA
A:Status: preliminary
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26630; NID:q405702; PIDN:CAA81383.1; PID:q405703
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
|||||||
Db 28 PKLLIYWA 35

RESULT 4

Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37529
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IgD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37529
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26631; NID:q405700; PIDN:CAA81384.1; PID:q405701
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
|||||||
Db 28 PKLLIYWA 35

RESULT 5

Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37535
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IgD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37535
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26625; NID:q405712; PIDN:CAA81378.1; PID:q405713
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
|||||||
Db 28 PKLLIYWA 35

RESULT 6

Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37534
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IgD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37534
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26626; NID:q405710; PIDN:CAA81379.1; PID:q405711
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
|||||||
Db 28 PKLLIYWA 35

RESULT 7

Ig kappa chain V region (V-kappa 4) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37532
R:Klein, U.; Kuipers, R.; Rajewsky, K.
submitted to the EMBL Data Library, September 1993
A:Description: Human IGM(+)IgD(+) cells, the major B cell subset in the peripheral blood
A:Reference number: S37501
A:Accession: S37532
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-92 <KLE>
A:Cross-references: EMBL:Z26628; NID:q405706; PIDN:CAA81381.1; PID:q405707
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.075;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
|||||||
Db 28 PKLLIYWA 35

RESULT 8

G38601

Ig kappa chain V region (4A9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Aug-1991 #sequence_revision 30-Aug-1991 #text_change 23-Jul-1999
C:Accession: G38601
R:Goshorn, S.C.; Retzel, E.; Jemerson, R.
J. Biol. Chem. 266, 2134-2142, 1991
A:Title: Common structural features among monoclonal antibodies binding the same antigen
A:Reference number: A38601; MUID:91115823
A:Accession: G38601
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-96 <GOS>
A:Cross-references: GB:M57984; NID:9196414; PIDN:AAA63365.1; PID:9196415
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 46; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.078;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLIYMA 8
|||||||
DB 32 PKLIYMA 39

RESULT 9
Ig kappa chain V region (M-1406) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 21-Jan-2000
C:Accession: S19974
R:Wissenhorn, W.; Riethmuller, G.; Weiss, E.M.; Rieber, E.P.
Submitted to the EMBL Data Library, March 1992
A:Description: Structural characterization of CD4 mAb.
A:Reference number: S19974
A:Accession: S19974
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-98 <MET>
A:Cross-references: EMBL:X65096; NID:952294; PIDN:CAA46224.1; PID:952295
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-80/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.08;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLIYMA 8
|||||||
DB 34 PKLIYMA 41

RESULT 10
Ig light chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1998 #sequence_revision 19-Mar-1998 #text_change 20-Jun-2000
C:Accession: S26337; S78449
R:Stark, S.E.; Caton, A.J.
J. Exp. Med. 174, 613-624, 1991
A:Title: Antibodies that are specific for a single amino acid interchange in a protein
A:Reference number: S26309; MUID:91341421
A:Accession: S26337
A:Molecule type: mRNA
A:Residues: 1-101 <STA>
A:Cross-references: EMBL:X59193
R:Caton, A.J.
submitted to the EMBL Data Library, April 1991
A:Reference number: S78447
A:Accession: S78449

A:Molecule type: mRNA
A:Residues: 1-60, 'T', '62-91', 'S', '93-101 <CAT>
A:Cross-references: EMBL:X59193; NID:952323; PIDN:CAA41903.1; PID:91334067
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:8-88/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLIYMA 8
|||||||
DB 42 PKLIYMA 49

RESULT 11
Ig light chain V region (clone 202.9) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PH1046
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1046
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-101 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:16-96/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLIYMA 8
|||||||
DB 50 PKLIYMA 57

RESULT 12
PH0869
Ig kappa chain V region (anti-DNA, H2F) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 21-Jan-2000
C:Accession: PH0869
R:Manheimer-Lory, A.; Katz, J.B.; Pillinger, M.; Ghossein, C.; Smith, A.; Diamond, B.
J. Exp. Med. 174, 1639-1652, 1991
A:Title: Molecular characteristics of antibodies bearing an anti-DNA-associated idiot,
A:Reference number: PH0862; MUID:92078875
A:Accession: PH0869
A:Molecule type: DNA
A:Residues: 1-101 <MAN>
A>Note: residues 28-33 were obtained from Figure 4
C:Comment: This antibody is produced by Epstein-Barr virus-transformed B cell that be
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-23/Region: framework 1
F:16-96/Domain: immunoglobulin homology <IMM>
F:24-40/Region: complementarity-determining 1
F:41-55/Region: framework 2
F:56-62/Region: complementarity-determining 2
F:63-94/Region: framework 3
F:95-101/Region: complementarity-determining 3

Query Match 100.0%; Score 46; DB 2; Length 101;

Best Local Similarity 100.0%; Pred. No. 0.082;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLLIYWA 8
|||||||
Db 50 PKLLIYWA 57

RESULT 13
PHI050
Ig light chain V region (clone 111-cl) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PHI050
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B
A:Reference number: PH0971; MUID:92381444
A:Accession: PHI050
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-103 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin
F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
|||||||
Db 50 PKLLIYWA 57

RESULT 14
PHI051
Ig light chain V region (clone 165.3) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PHI051
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective B
A:Reference number: PH0971; MUID:92381444
A:Accession: PHI051
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-103 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin
F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
|||||||
Db 50 PKLLIYWA 57

RESULT 15
PHI052
Ig light chain V region (clone 165.5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jan-2000
C:Accession: PHI052
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marion, T.N.

J. Exp. Med. 176, 761-779, 1992
A:Title: Both IgM and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PHI052
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-103 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: Immunoglobulin
F:16-96/Domain: Immunoglobulin homology <IMM>

Query Match 100.0%; Score 46; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
|||||||
Db 50 PKLLIYWA 57

Search completed: July 15, 2002, 13:01:01
Job time: 467 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:23:32 ; Search time 44.9 Seconds

(without alignments)
6.899 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46

Sequence: 1 PKLLIYWA 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	46	100.0	114 1 KVA4_HUMAN	P01625 homo sapien
2	46	100.0	121 1 KVA0_HUMAN	P06312 homo sapien
3	46	100.0	133 1 KVA8_HUMAN	P06313 homo sapien
4	46	100.0	134 1 KVA6_HUMAN	P06314 homo sapien
5	35	76.1	151 1 YCRC_RACCU	P42401 bacillus su
6	35	76.1	283 1 YORS_ADECI	P20747 avian adeno
7	35	76.1	705 1 PURL_PYROH	O59621 pyrococcus
8	34	73.9	146 1 AP4A_HUMAN	P50580 homo sapien
9	34	73.9	146 1 AP4A_MOUSE	P50584 sus scrofa
10	34	73.9	146 1 AP4A_PIG	P49792 homo sapien
11	34	73.9	3224 1 RBP2_HUMAN	P01605 homo sapien
12	33	71.7	108 1 KVM1_HUMAN	P01650 mus musculu
13	33	71.7	108 1 KVS0_MOUSE	P01651 mus musculu
14	33	71.7	108 1 KVS8_MOUSE	P01652 mus musculu
15	33	71.7	108 1 KVS5_MOUSE	P01653 mus musculu
16	33	71.7	108 1 KVS7_MOUSE	P01657 mus musculu
17	33	71.7	111 1 KVS3_MOUSE	P01662 mus musculu
18	33	71.7	111 1 KVS3_MOUSE	P01663 mus musculu
19	33	71.7	111 1 KVS3_MOUSE	P01667 mus musculu
20	33	71.7	111 1 KVS3_MOUSE	P01670 mus musculu
21	33	71.7	111 1 KVS3_MOUSE	P01671 mus musculu
22	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
23	33	71.7	111 1 KVS3_MOUSE	P01673 mus musculu
24	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
25	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
26	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
27	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
28	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
29	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
30	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
31	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
32	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu
33	33	71.7	111 1 KVS3_MOUSE	P01672 mus musculu

34	32	69.6	108 1 KV07_RABIT	P01688 oryctolagus
35	32	69.6	108 1 KV08_RABIT	P01689 oryctolagus
36	32	69.6	108 1 KV10_HUMAN	P01693 homo sapien
37	32	69.6	108 1 KV11_HUMAN	P01694 homo sapien
38	32	69.6	108 1 KV10_HUMAN	P01697 homo sapien
39	32	69.6	108 1 KV15_HUMAN	P01611 homo sapien
40	32	69.6	109 1 KV14_RABIT	P01695 oryctolagus
41	32	69.6	110 1 KV13_RABIT	P01694 oryctolagus
42	32	69.6	110 1 KV15_RABIT	P01696 oryctolagus
43	32	69.6	110 1 KV3P_MOUSE	P01668 mus musculu
44	32	69.6	110 1 KV12_RABIT	P01693 oryctolagus
45	32	69.6	111 1 KV3A_MOUSE	P01654 mus musculu

ALIGNMENTS

```
RESULT 1
KV4A_HUMAN 1
ID KV4A_HUMAN STANDARD: PRT: 114 AA.
AC P01625:
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1986 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region Len.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76004342; PubMed=50995;
RA Schneider M., Hilschmann N.;
RT "The primary structure of a monoclonic immunoglobulin-L-chain of
RT subgroup IV of the kappa type (Bence-Jones protein Len).";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1973).
RN [2]
RP REVISION TO 9.
RA Salomon A.;
RL Submitted (AUG-1996) to the SWISS-PROT data bank.
CC -I- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -I- MISCELLANEOUS: THIS IS A BENICE-JONES PROTEIN.
DR PIR: A01903; K4HJLN.
DR HSSP: P01789; IMCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 40 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 3 55 FRAMEWORK-2.
FT DOMAIN 4 62 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 5 94 FRAMEWORK-3.
FT DOMAIN 6 101 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 7 102 113 FRAMEWORK-4.
FT DISULEID 23 94 BY SIMILARITY.
FT NON_TER 114 114
SQ SEQUENCE 114 AA, 12640 MW; 0647F1D17F236485 CRC64;
```

Query Match 100.0%; Score 46; DB 1; Length 114;
Best local Similarity 100.0%; Pred. No. 0.03;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
Db 50 PKLLIYWA 57

RESULT 2
KV40_HUMAN STANDARD: PRT: 121 AA.
ID KV40_HUMAN

```

AC P06312;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=86041853; PubMed=2997712;
RA Klobbeck H.G., Bornkamm G.W., Combratato G., Mocikat R., Pohlentz H.D.,
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene.";
RL Nucleic Acids Res. 13:6515-6529(1985).
CC -1- MISCELLANEOUS: THERE IS ONLY ONE IG KAPPA V-IV GENE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 200023; CAA77318.1; -
DR PIR; A01902; KAHUJ.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 >121 IG KAPPA CHAIN V-IV REGION.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13380 MW; 9586AD418D33974 CRC64;

Query Match 100.0%; Score 46; DB 1; Length 121;
Best Local Similarity 100.0%; Pred. NO. 0.032;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
DB 70 PKLLIYMA 77

RESULT 3
KV4B-HUMAN STANDARD; PRT; 133 AA.
ID KV4B-HUMAN
AC P06313;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region JI precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=86041853; PubMed=2997712;
RA Klobbeck H.G., Bornkamm G.W., Combratato G., Mocikat R., Pohlentz H.D.,

```

```

RA Zachau H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene.";
RL Nucleic Acids Res. 13:6515-6529(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 200022; CAA77317.1; -
DR PIR; A01904; KAHUJI.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 133 IG KAPPA CHAIN V-IV REGION JI.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 122 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AR4 CRC64;

Query Match 100.0%; Score 46; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. NO. 0.035;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
DB 70 PKLLIYMA 77

RESULT 4
KV4C-HUMAN STANDARD; PRT; 134 AA.
ID KV4C-HUMAN
AC P06314;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region B17 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11;
RP SEQUENCE FROM N.A.
RX MEDLINE=86041854; PubMed=2997713;
RA Marsh P., Mills F., Gould H.;
RT "Detection of a unique human V kappa IV germline gene by a cloned
RT cDNA probe.";
RL Nucleic Acids Res. 13:6531-6544(1985).
RN 12;
RP REVISION TO 76.
RA Marsh P.;
RT Submitted (OCT-1986) to the EMBL/Genbank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: X02990; CAA26733.1; -
DR PIR: A01905; K4HUI7.
DR HSSP: P01789; IMCP.
DR InterPro: IPR003006; 19_MHC.
DR InterPro: IPR003596; 19_V.
DR Pfam: PF000047; 1g.1.
DR SMART: SM00406; IGV.1.
KM Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 1 134 IG KAPPA CHAIN V-IV REGION B17.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 121 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 122 133 FRAMEWORK-4.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 134 134
SQ SEQUENCE 134 AA: 14966 MW: 6413A22FD0738832 CRC64;

Query Match 100.0%; Score 46; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRLIYWA 8
| | | | | | | |
Db 70 PRLIYWA 77

RESULT 5
YKCSU
ID YKCSU STANDARD; PRT; 151 AA.
AC P42401;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 17.0 kDa protein in COM 5' region (ORF3).
GN YKCS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124189; PubMed=8969502;
RA Yamane K., Kumano M., Kurita K.;
RT "The 25 degrees-36 degrees region of the Bacillus subtilis
chromosome: determination of the sequence of a 146 kb segment and
RT identification of 113 genes.";
RL Microbiology 142:3047-3056(1996).
RN
RP SEQUENCE OF 43-151 FROM N.A.
RC STRAIN=168;
RX MEDLINE=95219080; PubMed=7704255;
RA Fujishima Y., Yamane K.;
RT of stfA of the Bacillus subtilis chromosome.";
RL Microbiology 141:277-279(1995).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

CC -----
DR EMBL: D50453; BAA08973.1; -
DR DR D30762; BAA06427.1; -
DR EMBL: Z99105; CAB12133.1; -
DR Subtilist; BG11179; ykcs.
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 14 34 POTENTIAL.
FT TRANSMEM 45 65 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
SQ SEQUENCE 151 AA: 16983 MW: 55395791ACDD3E0D CRC64;

Query Match 76.1%; Score 35; DB 1; Length 151;
Best Local Similarity 85.7%; Pred. No. 5.5;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRLIYW 7
| | | | | | |
Db 26 PRLIYW 32

RESULT 6
YOR5_ADEG1
ID YOR5_ADEG1 STANDARD; PRT; 283 AA.
AC P20747;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 31.5 kDa protein (ORF 5) (ORF10).
OS Avian adenovirus gal1 (strain Phelps) (Fowl adenovirus 1) (CELO).
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
OX NCBI_TaxID=10553;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90251474; PubMed=2160072;
RA Akopian T.A., Kruglyak V.A., Rivkina M.B., Naroditsky B.S.,
RA Tikhonenko T.I.;
RT "Sequence of an avian adenovirus (CELO) DNA fragment (0-11.2%).";
RL Nucleic Acids Res. 18:2825-2825(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96186720; PubMed=8627769;
RA Chiocca S., Kurzbauer R., Schaffner G., Baker A., Mautner V.,
RA Cotten M.;
RT "The complete DNA sequence and genomic organization of the avian
RT adenovirus CELO.";
RL J. Virol. 70:2939-2949(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
DR EMBL: X17217; CAA5087.1; -
DR DR EMBL: U46933; AAC54931.1; -
DR PIR: S10005; S10005.
KM Hypothetical protein.
SQ SEQUENCE 283 AA: 31487 MW: 6018412DA598183D CRC64;

Query Match 76.1%; Score 35; DB 1; Length 283;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PRLIYWA 8
| | | | | | |
Db 169 PRLIISWA 176

RESULT 7

PURL_PYRO
ID PURL_PYROH STANDARD: PRT: 705 AA.
AC 059621;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphoribosylformylglycinamide synthase II (EC 6.3.5.3) (FGAM synthase II).
DE PURL OR PH1953.
GN Pyrococcus horikoshii.
OS Pyrococcus horikoshii.
OC Archaea: Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohtuku Y., Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Ogunchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K., Masuchi Y., Shizuya H., Kikuchi H.;
RA "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:35-76(1998).
CC -1- CATALYTIC ACTIVITY: ATP + 5'-phosphoribosylformylglycinamide + L-glutamine + H(2)O = ADP + phosphate + 5'-phosphoribosylglycinamide + L-glutamate.
CC -1- PATHWAY: DE NOVO PURINE BIOSYNTHESIS, FOURTH STEP.
CC -1- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURO AND PURL.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE FGAMS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000007; BAA31080.1; -
DR InterPro: IPR000728; AIRS-related.
DR Pfam: PF00586; AIRS; 2.
DR Pfam: PF02769; AIRS_C; 2.
DR Purine biosynthesis; Ligase: ATP-binding; Complete proteome.
FT NP_BIND 89 100 ATP (POTENTIAL).
SQ SEQUENCE 705 AA; 78540 MM; 57550733MAADDE29 CRC64;

Query Match 76.1%; Score 35; DB 1; Length 705;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 PKLLTYW 7
DB 334 PKFVYW 340

RESULT 8
AP4A_HUMAN
ID AP4A_HUMAN STANDARD: PRT: 146 AA.
AC P50583;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bis(5'-nucleosyl)-tetraphosphatase (Asymmetrical) (EC 3.6.1.17)
DE (Diadenosine 5',5''-P1,P4-tetraphosphate asymmetrical hydrolase)
DE (Diadenosine tetraphosphatase) (AP4A hydrolase) (AP4ASE).
GN NUDT2 OR APAH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver and Spleen;
RX MEDLINE=96067583; PubMed=7487923;
RA Thorne N.M.H., Hankin S., Wilkinson M.C., Nunez C., Barracough R., McLennan A.G.;
RA "Human diadenosine 5',5''-P1,P4-tetraphosphate pyrophosphohydrolase is a member of the Nudt family of nucleotide pyrophosphatases.";
RL Biochem. J. 311:717-721(1995).
RN [2]
RP ACETYLATION.
RA McLennan A.G.;
RL Submitted (JAN-1997) to the SWISS-PROT data bank.
CC -1- FUNCTION: ASYMMETRICALLY HYDROLYZES AP4A TO YIELD AMP AND ATP.
CC -1- CATALYTIC ACTIVITY: P(1),P(4)-bis(5'-adenosyl)tetraphosphate + H(2)O = ATP + AMP.
CC -1- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U30313; AAC50277.1; -
DR MIM: 602852; -
DR InterPro: IPR000086; NUDIX_hydrolase.
DR Pfam: PF00293; NUDIX; 1.
DR PRINTS: PR00502; NUDIXFAMILY.
DR PROSITE: PS00893; NUDIX; 1.
KW Hydrolase; Acetylation.
FT INT_MET 0 0
FT DOMAIN 42 63 NUDIX BOX.
FT MOD_RES 1 1 ACETYLATION.
SQ SEQUENCE 146 AA; 16698 MM; 7C4D1B42CA487AB8 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 146;
Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PKLLTYW 7
DB 87 PKFVYW 93

RESULT 9
AP4A_MOUSE
ID AP4A_MOUSE STANDARD: PRT: 146 AA.
AC P56380;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bis(5'-nucleosyl)-tetraphosphatase (Asymmetrical) (EC 3.6.1.17)
DE (Diadenosine 5',5''-P1,P4-tetraphosphate asymmetrical hydrolase)
DE (Diadenosine tetraphosphatase) (AP4A hydrolase) (AP4ASE).
GN NUDT2 OR APAH1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Merra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M., Schellenberg K., Stepien M., Tan F., Underwood K., Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R., Waterston R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ASYMMETRICALLY HYDROLYZES AP4A TO YIELD AMP AND ATP (BY SIMILARITY).
CC

```
CC -1- CATALYTIC ACTIVITY: P(1),P(4)-bis(5'-adenosyl)tetrphosphate +
CC H(2)O -> ATP + AMP.
CC -1- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AA125383; -; NOT_ANNOTATED_CDS.
CC EMBL; AA222223; -; NOT_ANNOTATED_CDS.
CC EMBL; AA222233; -; NOT_ANNOTATED_CDS.
CC InterPro; IPR000086; NUDIX_hydrolase.
CC Pfam; PF00293; NUDIX; 1.
CC PRINTS; PR00502; NUDIXFAMILY.
CC PROSITE; PS00893; NUDIX; 1.
CC Hydrolase; Acetylation.
CC INIT_MET 0 BY SIMILARITY.
CC DOMAIN 42 63 NUDIX BOX.
CC MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
CC SEQUENCE 146 AA; 16943 MW; 333ABEFAD2E7671 CRC64;
SQ
Query Match 73.9%; Score 34; DB 1; Length 146;
Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PKLIYW 7
11:111
Db 87 PKTVIYW 93
RESULT 10
AP4A_PIG STANDARD; PRT; 146 AA.
AC P50584;
ID AP4A_PIG
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bis(5'-nucleosyl)-tetrphosphate (Asymmetrical) (EC 3.6.1.17)
DE (Diadenosine 5',5'-P1,P4-tetrphosphate asymmetrical hydrolase)
DE (Diadenosine tetraphosphatase) (AP4A hydrolase) (AP4AASE).
GN NUDP2 OR APAH1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RA Hanlin S., Winteroe A.R., McLennan A.G.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ASYMMETRICALLY HYDROLYSES AP4A TO YIELD AMP AND ATP.
CC -1- CATALYTIC ACTIVITY: P(1),P(4)-bis(5'-adenosyl)tetrphosphate +
CC H(2)O -> ATP + AMP.
CC -1- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U38619; AAB61380.1; -.
CC InterPro; IPR000086; NUDIX_hydrolase.
CC Pfam; PF00293; NUDIX; 1.
CC PRINTS; PR00502; NUDIXFAMILY.
CC PROSITE; PS00893; NUDIX; 1.
DR PROSITE; PS00893; NUDIX; 1.
```

```
KW Hydrolase; Acetylation.
FT INIT_MET 0 BY SIMILARITY.
FT DOMAIN 42 63 NUDIX BOX.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 146 AA; 16704 MW; 7AFEC59CBB4379 CRC64;
Query Match 73.9%; Score 34; DB 1; Length 146;
Best Local Similarity 71.4%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 PKLIYW 7
11:111
Db 87 PKTVIYW 93
RESULT 11
RBP2_HUMAN STANDARD; PRT; 3224 AA.
AC P49792; O15280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ran-binding protein 2 (RanBP2) (Nuclear pore complex protein Nup358)
DE (Nucleoporin Nup358) (358 kDa nucleoporin) (P270).
GN RanBP2 OR NUP358.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95294031; PubMed=7775481;
RA Wu J., Matunis M.J., Kraemer D., Blobel G., Coutavas E.;
RT "Nup358, a cytoplasmically exposed nucleoporin with peptide repeats,
RT Ran-GTP binding sites, zinc fingers, a cyclophilin A homologous
RT domain, and a leucine-rich region."
RL J. Biol. Chem. 270:14209-14213(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood.
RX MEDLINE=95327194; PubMed=7603572;
RA Yokoyama N., Hayashi N., Seki T., Hayashida T.,
RA Kuna K.I., Miyata T., Fukui M., Nishimoto T., Pante N., Aebl U.;
RT "A giant nucleopore protein that binds Ran/TC4."
RL Nature 376:184-188(1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.96 ANGSTROMS) OF 1171-1304.
RX MEDLINE=99176415; PubMed=10078529;
RA Vetter I.R., Nowak C., Nishimoto T., Kuhlmann J., Wittinghofer A.;
RT "Structure of a Ran-binding domain complexed with Ran bound to a GTP
RT analogue: implications for nuclear transport."
RL Nature 398:39-46(1999).
CC -1- FUNCTION: INVOLVED IN TRANSPORT FACTOR (RAN-GTP, KARYOPHEIN) -
CC MEDIATED PROTEIN IMPORT VIA THE F-G REPEAT-CONTAINING DOMAIN WHICH
CC ACTS AS A DOCKING SITE FOR SUBSTRATES. COULD ALSO HAVE ISOMERASE
CC OR CHAPERONE ACTIVITY AND MAY BIND RNA OR DNA. COMPONENT OF THE
CC NUCLEAR EXPORT PATHWAY. SPECIFIC DOCKING SITE FOR THE NUCLEAR
CC EXPORT FACTOR EXPORTIN-1.
CC -1- SUBUNIT: FORMS A TIGHT COMPLEX IN ASSOCIATION WITH RANBP1 AND THE
CC UBIQUITIN-CONJUGATING ENZYME E2 (UBC9) (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX. CYTOPLASMIC FILAMENTS.
CC -1- DOMAIN: CONTAINS MANY X-X-F-G AND X-F-X-F-G REPEATS.
CC -1- SIMILARITY: CONTAINS 8 RANBP2-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 4 RANBP2-TYPE ZINC FINGERS.
CC -1- SIMILARITY: CONTAINS 1 CYCLOPHILIN-LIKE PHASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```

CC or send an email to license@slb.ch).
CC -----
DR EMBL: L41840; AAC41758.1;
DR EMBL: D42063; BAA07662.1;
DR PDB: 1RRP; 1R-MAY-99.
DR MIM: 601181;
DR InterPro: IPR002130; CSA_PPIase.
DR InterPro: IPR000697; RanBP1_MASP.
DR InterPro: IPR000156; Ran_BP1.
DR InterPro: IPR001440; TPR.
DR InterPro: IPR001876; Znf-RanBP.
DR Pfam: PF00160; pro_isomerase; 1.
DR Pfam: PF00638; Ran_BP1; 4.
DR Pfam: PF00515; TPR; 1.
DR Pfam: PF00641; zf-RanBP; 8.
DR PRINTS: PR00153; CSAPISMASE.
DR SMART: SM00160; RanBP; 4.
DR SMART: SM00547; ZNF_RBP; 8.
DR PROSITE: PS00170; CSA_PPIASE_1; 1.
DR PROSITE: PS00772; CSA_PPIASE_2; 1.
DR PROSITE: PS01358; ZF_RANBP2_1; 8.
DR PROSITE: PS01099; ZF_RANBP2_2; 8.
KW Nuclear protein; Transport; Repeat; Zinc-finger; Isomerase; Rotamase;
KW 3D-structure.
FT DOMAIN 1172 1301 RANBP1-LIKE 1.
FT ZN_FING 1351 1381 RANBP2-TYPE 1.
FT ZN_FING 1415 1444 RANBP2-TYPE 2.
FT ZN_FING 1479 1508 RANBP2-TYPE 3.
FT ZN_FING 1543 1572 RANBP2-TYPE 4.
FT ZN_FING 1606 1635 RANBP2-TYPE 5.
FT ZN_FING 1665 1694 RANBP2-TYPE 6.
FT ZN_FING 1724 1753 RANBP2-TYPE 7.
FT ZN_FING 1781 1810 RANBP2-TYPE 8.
FT DOMAIN 2013 2142 RANBP1-LIKE 2.
FT DOMAIN 2310 2439 RANBP1-LIKE 3.
FT DOMAIN 2912 3040 RANBP1-LIKE 4.
FT DOMAIN 3067 3223 PPIASE, CYCLOPHILIN-TYPE.
FT CONFLICT 777 777 R -> H (IN REF. 2).
FT CONFLICT 784 784 R -> K (IN REF. 2).
SQ SEQUENCE 3224 AA; 358214 MW; 54E78412C96A3C63 CRC64;

Query Match 73.9%; Score 34; DB 1; Length 3224;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
Db 572 PALIYWA 579

RESULT 12
KVIM_HUMAN STANDARD; PRT; 108 AA.
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; Pubmed=824717;
RA Capra J D., Klapper D G.;
RT "Complete amino acid sequence of the variable domains of two human
RT IgM anti-gamma globulins (Lay/Pom) with shared idiotypic
RT specificities.*";
RL Scand. J. Immunol. 5:677-684(1976).
CC -I- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,

```

```

CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR: A01871; KIHUX.
DR HSSP: P01607; IRET.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 107 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;

Query Match 71.7%; Score 33; DB 1; Length 108;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
Db 44 PKLIYWA 51

RESULT 13
KV5Q_MOUSE STANDARD; PRT; 108 AA.
ID KV5Q_MOUSE
AC P01650;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region UPC 61.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; Pubmed=109517;
RA Vrana M., Rudikoff S., Potter M.;
RT "The structural basis of a hapten-inhibitable kappa-chain idiotype.*";
RL J. Immunol. 122:1905-1910(1979).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: A01929; KVM561.
DR HSSP: P80362; IWTI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11809 MW; FAE4DA36076F2AFE CRC64;

Query Match 71.7%; Score 33; DB 1; Length 108;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```


QY 1 PKLIIYWA 8
 |||||
 DB 44 PKLIIYGA 51

```

RESULT 14
KV5R_MOUSE STANDARD: PRT: 108 AA.
ID KV5R_MOUSE
AC P01651:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region EPC 109.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=79195288; PubMed=109517;
RA Vrana M., Rudikoff S., Potter M.;
RT "The structural basis of a hapten-inhibitable kappa-chain idiotype.";
RL J. Immunol. 122:1905-1910(1979).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: B92808; KWS09.
DR HSSP: P01607; IREI.
DR InterPro: IPR003306; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGv; 1.
KM Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11876 MW; 35C116BD60F79310 CRC64;

Query Match 71.7%; Score 33; DB 1; Length 108;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKLIIYWA 8
  |||||
DB 44 PKLIIYGA 51

RESULT 15
KV5S_MOUSE STANDARD: PRT: 108 AA.
ID KV5S_MOUSE
AC P01652:
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region J606.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Stankard J., Paul L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
binding proteins.";
RL J. Immunol. 128:302-307(1982).
```

```

CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM MYELOMA PROTEINS THAT
CC BIND BETA(2-1)-FRUCTOFURANOSYL MOIETIES (INULIN).
DR PIR: A92811; KWS06.
DR HSSP: P01607; IREI.
DR InterPro: IPR003306; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGv; 1.
KM Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 34 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49 FRAMEWORK-2.
FT DOMAIN 50 56 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 57 88 FRAMEWORK-3.
FT DOMAIN 89 97 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 98 108 FRAMEWORK-4.
FT DISULFID 23 88 BY SIMILARITY.
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11810 MW; 8DE4DD31076F2AFB CRC64;
```

```

Query Match 71.7%; Score 33; DB 1; Length 108;
Best Local Similarity 87.5%; Pred. No. 9.6;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKLIIYWA 8
  |||||
DB 44 PKLIIYGA 51
```

Search completed: July 15, 2002, 13:23:32
 Job time: 1448 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:22:42 ; Search time 172.49 Seconds
(without alignments)
8.023 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46

Sequence: 1 PKLIIYWA 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertbrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.cheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	130	4 Q9NP29	Q9NP29 homo sapien
2	41	89.1	2703	10 Q9MAZ4	Q9MAZ4 arabidopsis
3	41	89.1	2810	10 Q9FKS4	Q9FKS4 arabidopsis
4	37	80.4	99	11 Q9JL74	Q9JL74 mus musculu
5	37	80.4	112	9 Q38357	Q38357 lactococcus
6	37	80.4	227	10 Q23209	Q23209 arabidopsis
7	37	80.4	375	16 Q92KS3	Q92KS3 rhizobium m
8	36	78.3	291	11 Q35558	Q35558 mus musculu
9	35	76.1	288	5 Q9VRI6	Q9VRI6 drosophila
10	35	76.1	329	6 Q97885	Q97885 equus cabal
11	35	76.1	362	17 Q9V2I6	Q9V2I6 pyrococcus
12	34	73.9	147	11 Q9D6V2	Q9D6V2 mus musculu
13	34	73.9	147	11 Q9D6U6	Q9D6U6 mus musculu
14	34	73.9	202	12 Q9TDM3	Q9TDM3 avian adeno
15	34	73.9	303	12 Q64777	Q64777 avian adeno
16	34	73.9	367	10 Q9STR5	Q9STR5 arabidopsis

17	34	73.9	401	10 Q64859	Q64859 arabidopsis
18	34	73.9	402	10 Q944H9	Q944H9 arabidopsis
19	34	73.9	438	2 Q87856	Q87856 streptomyces
20	34	73.9	496	17 Q59477	Q59477 pyrococcus
21	34	73.9	556	12 Q36369	Q36369 alcelaphine
22	34	73.9	639	17 Q59084	Q59084 pyrococcus
23	34	73.9	706	2 Q88090	Q88090 enterococcus
24	34	73.9	815	16 Q9KLG3	Q9KLG3 vibrio chol
25	34	73.9	904	4 Q9H0B2	Q9H0B2 homo sapien
26	34	73.9	1666	5 Q917P4	Q917P4 drosophila
27	34	73.9	1765	4 Q96656	Q96656 homo sapien
28	34	73.9	3053	11 Q9ER09	Q9ER09 mus musculu
29	34	72.8	438	10 Q94A82	Q94A82 arabidopsis
30	33	71.7	111	11 Q920E9	Q920E9 mus musculu
31	33	71.7	244	10 Q9L0W2	Q9L0W2 arabidopsis
32	33	71.7	256	16 Q67471	Q67471 aquifex aeo
33	33	71.7	519	16 Q9KTN8	Q9KTN8 vibrio chol
34	33	71.7	527	16 Q9KLD9	Q9KLD9 vibrio chol
35	33	71.7	687	3 Q74520	Q74520 schizosacch
36	33	71.7	817	5 Q9V2J5	Q9V2J5 drosophila
37	33	71.7	939	16 Q9PLB7	Q9PLB7 chlamydia m
38	33	71.7	956	16 Q84812	Q84812 chlamydia t
39	33	71.7	1483	3 Q96UR6	Q96UR6 diaportha a
40	32	69.6	103	11 Q9JL80	Q9JL80 mus musculu
41	32	69.6	107	4 Q96SA9	Q96SA9 homo sapien
42	32	69.6	110	16 Q84455	Q84455 chlamydia t
43	32	69.6	153	10 Q04520	Q04520 arabidopsis
44	32	69.6	183	4 Q9Y2V0	Q9Y2V0 homo sapien
45	32	69.6	217	17 Q27742	Q27742 methanother

ALIGNMENTS

RESULT 1
ID Q9NP29 PRELIMINARY: PRT; 130 AA.
AC Q9NP29;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MICROTUBULIN PROTEIN 2 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP MEDLINE=96354815; PubMed=8753791;
RX Osvath K.J., Xia S., Hirose H., Tilson M.D.;
RT "Two hypothetical proteins of human aortic adventitia, with Ig kappa,
RT collagenous, and aromatic-rich motifs.";
RL Biochem. Biophys. Res. Commun. 225:500-504(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97367690; PubMed=9224393;
RA Osvath K.J., Hirose H., Xia S., Chew D., Knoetgen J. III,
RA Tilson M.D.;
RT "Expression of two novel recombinant proteins from aortic adventitia
RT (kappa) sharing amino acid sequences with cytomegalovirus.";
RL J. Surg. Res. 69:277-282(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX Osvath K.J., Xia S., Hirose H., Tilson M.D.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF206020; AAF62402.1; -
DR HSSP: P80362; 1WTU.
DR InterPro: IPR003600; Ig-like.
DR SMART: SM00410; Ig-like; I.
FT NON_TER 130
SQ SEQUENCE 130 AA; 14128 MW; 51275185ACC6FA1E CRC64;

Query Match
Best Local Similarity 100.0%; Score 46; DB 4; Length 130;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 8
Db 70 PKLLIYW 77

RESULT 2
Q9MAZ4 PRELIMINARY; PRT; 2703 AA.

AC Q9MAZ4: 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE ATRAD3.
GN ATRAD3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;
RA Sugiyama H., Oguchi K., Tamura K., Takahashi H.;
RT "Arabidopsis thaliana mRNA for ATRAD3, complete cds."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB040133; BAA52828.1; -
DR InterPro: IPR003151; FAT.
DR InterPro: IPR003152; FATC.
DR InterPro: IPR00403; P13_P14_kinase.
DR Pfam: PF02259; FAT; 1.
DR Pfam: PF02260; FATC; 1.
DR Pfam: PF00454; P13_P14_kinase; 1.
DR SMART: SM00146; PI3K; 1.
DR PROSITE: PS50290; P13_4_KINASE_3; 1.
DR PROSITE: PS50290; P13_4_KINASE_3; 1.
SQ SEQUENCE 2703 AA; 302511 MW; B171F21EA12AD04 CRC64;

Query Match
Best Local Similarity 89.1%; Score 41; DB 10; Length 2703;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 7
Db 972 PKLLIYW 978

RESULT 3
Q9FKS4 PRELIMINARY; PRT; 2810 AA.

AC Q9FKS4: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN ATRAD3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=COLUMBIA;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
features of the regions of 1,381,565 bp covered by twenty one
physically assigned P1 and TAC clones.";
RL DNA Res., 5:131-145(1998).

DR EMBL: AB011477; BAB1344.1; -
DR InterPro: IPR003151; FAT.
DR InterPro: IPR003152; FATC.
DR InterPro: IPR002016; Peroxidase.
DR InterPro: IPR00403; P13_P14_kinase.
DR Pfam: PF02259; FAT; 1.
DR Pfam: PF02260; FATC; 1.
DR Pfam: PF00454; P13_P14_kinase; 1.
DR SMART: SM00146; PI3K; 1.
DR PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE: PS50290; P13_4_KINASE_3; 1.
DR PROSITE: PS50290; P13_4_KINASE_3; 1.
SQ SEQUENCE 2810 AA; 314057 MW; FB82AETCA60BD0E7 CRC64;

Query Match
Best Local Similarity 89.1%; Score 41; DB 10; Length 2810;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 7
Db 971 PKLLIYW 977

RESULT 4
Q9JL74 PRELIMINARY; PRT; 99 AA.

AC Q9JL74: 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN ANTI-MYOSIN IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C;
RA Melkiel S., Liao L., Cunningham M.W., Diamond B.;
RT "T-Cell-dependent antibody response to the dominant epitope of
streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive
with cardiac myosin.";
RL Infect. Immun. 68:5803-5808(2000).
DR EMBL: AF206032; AAF69330.1; -
DR HSSP: P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 99
SQ SEQUENCE 99 AA; 10939 MW; 3B25D0E784533324 CRC64;

Query Match
Best Local Similarity 80.4%; Score 37; DB 11; Length 99;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYW 8
Db 36 PKLLIYW 43

RESULT 5
Q38357 PRELIMINARY; PRT; 112 AA.

AC Q38357: 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
GN HYPOTHETICAL 13.2 KDA PROTEIN.

OS Lactococcus delbrueckii bacteriophage LL-H.
 OC Viruses.
 OX NCBI_TaxID=12348;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93267750; PubMed=8497043;
 RA Vasala A., Dupont L., Baumann M., Ritzenthaler P., Alatosassa T.;
 RT "Molecular comparison of the structural proteins encoding gene
 RT clusters of two related Lactobacillus delbrueckii bacteriophages.";
 RL J. Virol. 67:3061-3068(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94237431; PubMed=7514146;
 RA Mikkonen M., Vuorisalo J., Alatosassa T.;
 RT "Ribosome binding site consensus sequence of Lactobacillus delbrueckii
 RT subsp. lactis bacteriophage LL-H.";
 RL FEMS Microbiol. Lett. 116:315-320(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95129893; PubMed=7828907;
 RA Mikkonen M., Alatosassa T.;
 RT "Characterization of the genome region encoding structural proteins of
 RT Lactobacillus delbrueckii subsp. lactis bacteriophage LL-H.";
 RL Gene 151:53-59(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96064414; PubMed=8526515;
 RA Vasala A., Valkkila M., Caldentey J., Alatosassa T.;
 RT "Genetic and biochemical characterization of the Lactobacillus
 RT delbrueckii subsp. lactis bacteriophage LL-H LysIn.";
 RL Appl. Environ. Microbiol. 61:4004-4011(1995).
 DR EMBL: M96254; AAC00555.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 13203 MW; 81B615035BA7AFFE CRC64;

Query Match 80.4%; Score 37; DB 9; Length 112;
 Best Local Similarity 75.0%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PKLIYWA 8
 I I I I I I
 DB 22 PTLMTWA 29

RESULT 6
 O32209 PRELIMINARY; PRT; 227 AA.
 AC O32209;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 25.2 KDA PROTEIN.
 GN C7A10.640 OR A14G36720.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Terry N., Vos P., Heljnen L., Mewes H.W., Schueller C.,
 RA Chalvatiz N.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z99708; CAB16825.1; -
 DR EMBL: A161589; CAB80338.1; -
 DR InterPro: IPR004345; TB2_DPL_HVA22.
 DR Pfam: PF03134; TB2_DPL_HVA22; 1.
 KW Hypothetical protein.

SQ SEQUENCE 227 AA; 25231 MW; 14C52A404EE3E00 CRC64;
 Query Match 80.4%; Score 37; DB 10; Length 227;
 Best Local Similarity 85.7%; Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KLIYWA 8
 I I I I I I
 DB 67 KMLIYWA 73

RESULT 7
 O92KS3 PRELIMINARY; PRT; 375 AA.
 AC O92KS3;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL SIGNAL PEPTIDE PROTEIN SMC00400.
 GN SMC00400.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21368234; PubMed=11474104;
 RA Galibert F., Flan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Bailly-Hubler F., Barnett M.J., Becker A., Bolstad P., Botne G.,
 RA Boulry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
 RA Cowle A., Davis R.W., Dreano S., Felderspiel N.A., Fisher R.F.,
 RA Gloux S., Godrie T., Goffeau A., Gouzy J., Gurjal M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelie D., Purnelle B.,
 RA Ramsperger U., Surzycki R., Thebaud P., Vandenbol M.,
 RA Vorhoeiter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";
 RL Science 293:668-672(2001).
 DR EMBL: AL591783; CAC41757.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 375 AA; 39924 MW; 79592B7A7673B28 CRC64;

Query Match 80.4%; Score 37; DB 16; Length 375;
 Best Local Similarity 85.7%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 PKLIYWA 7
 I I I I I I
 DB 264 PKLIYWA 270

RESULT 8
 O35558 PRELIMINARY; PRT; 291 AA.
 AC O35558;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ERK2.
 GN MAPK1 OR ERK2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97407954; PubMed=9261178;
 RA Suglura N., Suga T., Ozeki Y., Maniwa G., Takishima K.;
 RT "The mouse extracellular signal-regulated kinase 2 gene: gene

RT structure and characterization of the promoter."
 CC J. Biol. Chem. 272:21575-21581(1997).
 -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; D87271; BAA22620.1; JOINED.
 DR EMBL; D87264; BAA22620.1; JOINED.
 DR EMBL; D87265; BAA22620.1; JOINED.
 DR EMBL; D87266; BAA22620.1; JOINED.
 DR EMBL; D87267; BAA22620.1; JOINED.
 DR EMBL; D87268; BAA22620.1; JOINED.
 DR EMBL; D87269; BAA22620.1; JOINED.
 DR HSRP; P27703; IERK.
 DR MGD; MG1:1346858; Mapk1.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003527; Map_kin.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS01351; MAPK; UNKNOWN; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR ATP-binding; Serine/threonine-protein kinase; Transferase.
 KW SEQUENCE 291 AA; 33601 MW; 7088216F35A8A260 CRC64;

Query Match 78.3%; Score 36; DB 11; Length 291;
 Best Local Similarity 85.7%; Pred. No. 43;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PKLIYW 7
 ||| |||
 Db 282 PKLIYW 288

RESULT 9
 O9VRL6 PRELIMINARY; PRT; 288 AA.
 ID O9VRL6
 AC O9VRL6
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE CG15214 PROTEIN.
 GN CG15214.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE=20196006; Pubmed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wotman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borokova D., Botchan M.R., Bouck J., Broksstein P., Brottler P.,
 Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodde K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Hartis N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ijzerman K.,
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Mishina N.V., Mohanty C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003566; AAF50779.1; -.
 DR FLYbase; FBgn0035606; CG15214.
 DR InterPro; IPR003598; Ig_c2.
 DR InterPro; IPR003006; Ig_MHC.
 DR Pfam; PF00047; Ig; 1.
 DR SMART; SM00408; IgC2; 1.
 KW Immunoglobulin domain.
 KW SEQUENCE 288 AA; 31235 MW; 8366F44C601617FE CRC64;

Query Match 76.1%; Score 35; DB 5; Length 288;
 Best Local Similarity 71.4%; Pred. No. 65;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 PKLIYW 7
 ||| |||
 Db 27 PKLIYW 33

RESULT 10
 ID O97885 PRELIMINARY; PRT; 329 AA.
 AC O97885;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE TYROSINASE-RELATED PROTEIN 1 (FRAGMENT).
 GN TRP1.
 OS Equus caballus (Horse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OC NCBI_Taxid=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKIN;
 RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
 RT "An equine sequence homologous to tyrosinase-related protein-1 (TRP1)
 mapped to chromosome 23/14-16."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF076781; AAC97109.1; -.
 DR InterPro; IPR002227; Tyrosinase.
 DR Pfam; PF00264; tyrosinase; 2.
 DR PRINTS; PR00092; TYROSINASE.
 DR PROSITE; PS00497; TYROSINASE_1; 1.
 FT NON_TER 1
 FT NON_TER 329
 FT SEQUENCE 329 AA; 37447 MW; 1258855FA453F864 CRC64;

Query Match 76.1%; Score 35; DB 6; Length 329;
 Best Local Similarity 50.0%; Pred. No. 74;
 Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PKLIYWA 8
 |||:|:|:
 Db 212 PKLIYWA 219

```

RESULT 11
O9V2L6      PRELIMINARY:      PRT:      362 AA.
ID O9V2L6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE DOLICHO-L-PHOSPHATE MANNOSYLTRANSFERASE.
GN PAB2307.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=23292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ORSAV;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248283; CAB48982.1;
DR InterPro; IPR001173; Glycos_transf_2.
DR Pfam; PF00535; Glycos_transf_2; 1.
KW Complete proteome.
SQ SEQUENCE 362 AA; 41505 MW; 50A99F7195F540DE CRC64;

```

```

Query Match      76.1%; Score 35; DB 17; Length 362;
Best Local Similarity 85.7%; Pred. No. 81;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 2 KLLIYWA 8
Db 323 QLLIYWA 329

```

```

RESULT 12
O9D6V2      PRELIMINARY:      PRT:      147 AA.
ID O9D6V2
AC O9D6V2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2310051L06RIK PROTEIN.
GN 2310051L06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";

```

```

RL Nature 409:685-690(2001).
DR EMBL; AK009933; BAB26592.1;
DR MGD; MGI:1913651; 2310051L06RIK.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR PRINTS; PR00502; NUDIXFAMILY.
DR PROSITE; PS00893; NUDIX; 1.
SQ SEQUENCE 147 AA; 16989 MW; A634D19F1E8CE34 CRC64;

```

```

Query Match      73.9%; Score 34; DB 11; Length 147;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 PKLIYW 7
Db 88 PKTIYW 94

```

```

RESULT 13
O9D2U6      PRELIMINARY:      PRT:      147 AA.
ID O9D2U6
AC O9D2U6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2310051L06RIK PROTEIN.
GN 2310051L06RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=CEREBELLUM;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";

```

```

RT Nature 409:685-690(2001).
DR EMBL; AK018771; BAB31399.1;
DR MGD; MGI:1913651; 2310051L06RIK.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR PRINTS; PR00502; NUDIXFAMILY.
DR PROSITE; PS00893; NUDIX; 1.
SQ SEQUENCE 147 AA; 16957 MW; APD3C5CF1E8CDA64 CRC64;

```

```

Query Match      73.9%; Score 34; DB 11; Length 147;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 PKLIYW 7
Db 88 PKTIYW 94

```

```

RESULT 14

```

Q91BM3
 ID Q91BM3 PRELIMINARY; PRT; 202 AA.
 AC Q91BM3;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE ORF18.
 OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
 ON NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PHELPS (ATCC VR-432);
 RX MEDLINE=96186720; PubMed=8627769;
 RA Chlocca S., Kurtzbaue R., Schaffner G., Baker A., Mautner V.,
 RA Corten M.;
 RT "The complete DNA sequence and genomic organization of the avian
 RT adenovirus CELO.";
 RL J. Virol. 70:2939-2949(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PHELPS (ATCC VR-432);
 RA Corten M.L.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U46933; AAC54925.1; -;
 SQ SEQUENCE 202 AA; 23210 MW; C0177BD547D2F67 CRC64;

Query Match 73.9%; Score 34; DB 12; Length 202;
 Best Local Similarity 57.1%; Pred. No. 72;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRLIYW 7
 DB 39 PRLIYW 45

RESULT 15
 ID 064777 PRELIMINARY; PRT; 303 AA.
 AC 064777;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ORF5 PROTEIN (FRAGMENT).
 OS Avian adenovirus gall (strain Phelps) (Fowl adenovirus 1) (CELO).
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Aviadenovirus.
 ON NCBI_TaxID=10553;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CELO;
 RA Akopian T.A., Kaverina E.N., Kruglyak V.A., Naroditsky B.S.,
 RA Tikhonenko T.T.;
 RT "Sequence of an Avian Adenovirus (CELO) DNA Fragment (11.2 - 19.2
 RT %)."
 RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z22864; CAA80481.1; -;
 FT NON_TER 303 303
 SQ SEQUENCE 303 AA; 34492 MW; 0A34B0F941D513C CRC64;

Query Match 73.9%; Score 34; DB 12; Length 303;
 Best Local Similarity 57.1%; Pred. No. 1.1e+02;
 Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRLIYW 7
 DB 39 PRLIYW 45

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:57:56 ; Search time 228.39 Seconds
(without alignments)
3.891 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46
Sequence: 1 PKLLIYWA 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A.GeneSeq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	82	19	AAW62807
2	46	100.0	84	17	AAW14491
3	46	100.0	84	17	AAW99878
4	46	100.0	94	19	AAW62808
5	46	100.0	100	19	AAW70948
6	46	100.0	106	19	AAW31647
7	46	100.0	107	13	AAW25723
8	46	100.0	107	14	AAW38599
9	46	100.0	107	19	AAW80817
10	46	100.0	107	19	AAW80814
11	46	100.0	107	19	AAW58499

12	46	100.0	107	19	AAW47087
13	46	100.0	107	20	AAW17955
14	46	100.0	107	21	AAW90375
15	46	100.0	107	21	AAW79425
16	46	100.0	107	22	AAW69665
17	46	100.0	107	22	AAW69666
18	46	100.0	107	22	AAW6226
19	46	100.0	110	19	AAW30012
20	46	100.0	110	19	AAW65520
21	46	100.0	110	19	AAW65517
22	46	100.0	111	15	AAW52059
23	46	100.0	112	13	AAW28807
24	46	100.0	112	13	AAW4932
25	46	100.0	112	15	AAW4933
26	46	100.0	113	15	AAW50313
27	46	100.0	113	15	AAW50314
28	46	100.0	113	15	AAW50322
29	46	100.0	113	15	AAW59513
30	46	100.0	113	15	AAW6964
31	46	100.0	113	15	AAW63770
32	46	100.0	113	15	AAW63769
33	46	100.0	113	17	AAW92215
34	46	100.0	113	17	AAW92217
35	46	100.0	113	18	AAW27695
36	46	100.0	113	20	AAW50145
37	46	100.0	113	20	AAW42267
38	46	100.0	113	20	AAW42268
39	46	100.0	113	20	AAW42269
40	46	100.0	113	20	AAW05759
41	46	100.0	113	20	AAW97177
42	46	100.0	113	21	AAW18855
43	46	100.0	113	21	AAW18857
44	46	100.0	113	21	AAW18861
45	46	100.0	113	21	AAW18863

ALIGNMENTS

AAW62807	RESULT 1
ID	AAW62807 standard; Peptide: 82 AA.
AC	AAW62807:
XX	
XX	
XX	23-SEP-1998 (first entry)
DE	Amino acid sequence of a human antibody fragment.
XX	
KW	Human; immunoglobulin; Ig; transgenic; non-human mammal;
KW	inactivated endogenous Ig locus; B-cell development;
KW	human heavy chain Ig locus; micro constant region; J-H; D-H; V-H gene;
KW	kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
KW	production; antibody.
OS	Homo sapiens.
XX	
PN	WO9824893-A2.
XX	
PD	11-JUN-1998.
XX	
PF	03-DEC-1997; 97WO-US23091.
XX	
PR	03-DEC-1996; 96US-0759620.
XX	
PA	(ABGE-) ABGENIX INC.
PI	Green L, Jakobovits A, Klapholz S, Kucherlapati R;
PI	Mendez M;
XX	
DR	WPI; 1998-333314/29.
XX	
PT	New transgenic non-human mammals - having an inactivated

Mouse J591 monoclonal
Human kappa 8 light
J591 monoclonal an
Tie2 receptor antia
Murine Fd138-80 an
Humanised Fd138-80
Monoclonal antibody
Light chain variab
Anti-DNA IL/1M (H2
H221 antibody VL r
Light chain variab
5A8 VL. Synthetic
Fc receptor humani
Mab 022 VK chain.
Humanised light ch
Humanised light ch
Mab NF52 light cha
Sequence of the ma
CC49 V-light. Mus
Humanised L chain
Murine L chain fro
LL2 Mab VK region.
Humanised LL2 Mab
Variable kappa cha
Antibody p19 human
Murine anti-TAG-72
Human REI antibody
Humanised CC49 VK
Anti-TAG-72 Mab CC
Monoclonal antibody
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc

PT Immunoglobulin locus and a near complete human immunoglobulin locus,
 PT used for production of human antibodies
 XX
 PS Disclosure: Page 78; 128pp; English.
 CC AAW62793-822 represent fragments of human antibodies produced by
 CC transgenic Xenomice, created using the method of the invention. The
 CC specification describes a transgenic non-human mammal which has genome
 CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
 CC locus, so that the mammal does not display normal B-cell development. The
 CC modified genome also has an inserted human heavy chain Ig locus in
 CC germline configuration, the human heavy chain Ig locus comprising a human
 CC micro constant region and regulatory and switch sequences, human J-H
 CC genes, human D-H genes, and human V-H genes and an inserted human kappa
 CC light chain Ig locus in germline configuration, the human kappa light
 CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
 CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
 CC are selected to restore normal B-cell development in the mammal. The
 CC transgenic animals have a near complete human Ig locus, including both a
 CC human heavy chain locus and a human kappa light chain locus. They can
 CC be used for the production of human antibodies when exposed to
 CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
 CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
 CC respectively.
 CC
 SQ Sequence 82 AA;

Query Match 100.0%; Score 46; DB 19; Length 82;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
 |||||
 Db 31 PKLLIYWA 38

RESULT 2
 AAW14491
 ID AAW14491 standard; Protein: 84 AA.
 XX
 AC AAW14491;
 XX
 DT 28-JAN-1997 (first entry)
 XX
 DE Monoclonal antibody D VK.
 XX
 KW heavy; light chain; monoclonal antibody; antigen 3; marker; melanoma;
 KW permanent human tumour cell line; tumour-associated antigen; epitope;
 KW gastrointestinal tumour; pancreatic carcinoma; diagnostic; therapeutic;
 KW antigen 11; Vibrio cholera; neuraminidase-resistant; ganglioside GD2.
 XX
 OS Synthetic.
 XX
 PN EP727436-A1.
 XX
 PD 21-AUG-1996.
 XX
 PF 21-MAR-1990; 90EP-0105322.
 XX
 PR 24-MAR-1989; 89DE-3909799.
 XX
 PA (BEHW) BEHRINGWERKE AG.
 XX
 PI Auerbach B, Bosslet K, Sedlacek H, Seemann G;
 XX
 DR WPI: 1996-372836/38.
 DR N-PSDB; AAT63508.
 XX
 PT Monoclonal antibody to tumour-associated antigen - useful as
 PT gastrointestinal tumour marker
 XX
 PS Disclosure: Page 14; 19pp; German.

XX
 CC AAW14490-91 are the heavy and light chains (respectively) of monoclonal
 CC antibody (Mab) D. Mab D recognises Vibrio cholera
 CC neuraminidase-resistant epitope of ganglioside GD2, from a human melanoma
 CC cell line. Mabs A, B and C (see AAW14484-89) are mentioned in the
 CC specification, but are not part of the claims. Mabs A and B recognise
 CC antigens 3 and 11 resp., of a permanent human tumour cell line. Mab C
 CC also recognises an epitope of a tumour-associated antigen. These antigens
 CC occur at high concns. in the serum of patients with gastrointestinal
 CC tumours, e.g. pancreatic carcinoma, and are thus useful as tumour markers
 CC for diagnostic or therapeutic purposes.
 CC
 SQ Sequence 84 AA;

Query Match 100.0%; Score 46; DB 17; Length 84;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLLIYWA 8
 |||||
 Db 22 PKLLIYWA 29

RESULT 3
 AAR9878
 ID AAR9878 standard; Protein: 84 AA.
 XX
 AC AAR9878;
 XX
 DT 28-JAN-1997 (first entry)
 XX
 DE Monoclonal antibody D VK.
 XX
 KW Monoclonal antibody; Mab; epitope; tumour-associated antigen;
 KW marker; antigen.
 XX
 OS Synthetic.
 XX
 PN EP727435-A1.
 XX
 PD 21-AUG-1996.
 XX
 PF 21-MAR-1990; 90EP-0105322.
 XX
 PR 24-MAR-1989; 89DE-3909799.
 XX
 PA (BEHW) BEHRINGWERKE AG.
 XX
 PI Auerbach B, Bosslet K, Sedlacek H, Seemann G;
 XX
 DR WPI: 1996-372835/38.
 DR N-PSDB; AAT36666.
 XX
 PT Monoclonal antibody to tumour-associated antigen - useful as
 PT gastrointestinal tumour marker
 XX
 PS Disclosure: Page 14; 19pp; German.

CC Mab C (AAT36659-T36660) is a monoclonal antibody that recognises an
 CC epitope of a tumour-associated antigen occurring at high concn. in
 CC the serum of patients with gastrointestinal tumours, e.g. pancreatic
 CC carcinoma, and is thus useful as a tumour marker for diagnostic or
 CC therapeutic purposes.
 CC Mabs A, B and D are mentioned in the specification, but are not
 CC part of the claims.
 CC Mab A (AAT36661-T36662) recognises antigen 3 of permanent human
 CC tumour cell line.
 CC Mab B (AAT36663-T36664) recognises antigen 11 of permanent human
 CC tumour cell line.
 CC Mab D (AAT36665-T36666) recognises a Vibrio cholera neuraminidase-
 CC resistant epitope of ganglioside GD2, from a human melanoma cell
 CC line.

XX Sequence 84 AA;
SQ

Query Match 100.0%; Score 46; DB 17; Length 84;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
|||||||
DB 22 PKLIYWA 29

RESULT 4

AAW62808
ID AAW62808 standard; Peptide: 94 AA.

XX
AC AAW62808;

XX
DT 23-SEP-1998 (first entry)

XX Amino acid sequence of a human antibody fragment.

DE Human; immunoglobulin; Ig; transgenic; non-human mammal;

XX Inactivated endogenous Ig locus; B-cell development;

KW human heavy chain Ig locus; micro constant region; J-H: D-H: V-H gene;

KW kappa light chain Ig locus; kappa constant region; J-kappa gene; V-kappa;
production; antibody.

XX Homo sapiens.

OS
XX MO9824893-A2.

PN 11-JUN-1998.

PD 03-DEC-1997; 97WO-US23091.

PF 03-DEC-1996; 96US-0759620.

PR (ABGE-) ABGENIX INC.

PA Green L, Yakobovits A, Klapholz S, Kucherlapati R;

PI Mendez M;

XX WPI: 1998-333314/29.

DR New transgenic non-human mammals - having an inactivated
PT immunoglobulin locus and a near complete human immunoglobulin locus,
PT used for production of human antibodies

XX Disclosure: Page 79; 128pp; English.

XX AAW62793-822 represent fragments of human antibodies produced by
CC transgenic Xenomice, created using the method of the invention. The
CC specification describes a transgenic non-human mammal which has genome
CC modifications that comprise an inactivated endogenous immunoglobulin (Ig)
CC locus, so that the mammal does not display normal B-cell development. The
CC modified genome also has an inserted human heavy chain Ig locus in
CC germline configuration, the human heavy chain Ig locus comprising a human
CC micro constant region and regulatory and switch sequences, human J-H
CC genes, human D-H genes, and human V-H genes and an inserted human kappa
CC light chain Ig locus in germline configuration, the human kappa light
CC chain Ig locus comprising a human kappa constant region, J-kappa genes,
CC and V-kappa genes, where the number of V-H and V-kappa genes inserted
CC are selected to restore normal B-cell development in the mammal. The
CC transgenic animals have a near complete human Ig locus, including both a
CC human heavy chain locus and a human kappa light chain locus. They can
CC be used for the production of human antibodies when exposed to
CC particular antigens e.g. when exposed to human IL-8, EGFR or TNF- alpha
CC the mice will produce antibodies to IL-8, EGFR or TNF- alpha
CC respectively.

XX Sequence 94 AA;
SQ

Query Match 100.0%; Score 46; DB 19; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.24; 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

QY 1 PKLIYWA 8
|||||||
DB 31 PKLIYWA 38

RESULT 5

AAW70948
ID AAW70948 standard; protein: 100 AA.

XX
AC AAW70948;

XX 14-OCT-1998 (first entry)

DE Light chain of monoclonal antibody RS-348.

XX Pathogenic virus; tropism: mucosa; CDR region; monoclonal antibody;

KW respiratory syncytial virus; RSV; VP6 protein; rota virus; RV;

KW viral infection; inhibit; fusion; protection; transcription;

KW antiviral agent; prophylaxis; diagnosis; infection; contamination.

XX Mus sp.

OS
XX FR2758331-A1.

PN 17-JUL-1998.

PD 14-JAN-1997; 97FR-0000300.

PF 14-JAN-1997; 97FR-0000300.

PR (UYBO-) UNIV BOURGOGNE.

PA Bourgeois C, Kohli E, Pothier P;

PI WPI: 1998-390320/34.

DR New peptide(s) recognising viral epitope with tropism to mucosa -
PT useful for, e.g. diagnosing, preventing and treating viral
PT infection(s)

XX Disclosure: Fig 2; 51pp; French.

XX The present sequence represents the light chain of monoclonal antibody
CC RS-348, which is directed against the respiratory syncytial virus (RSV).
CC The specification describes peptides which recognise, by
CC antigen-antibody type reactions, at least 1 epitope of a pathogenic virus
CC having tropism for the mucosa. AAW70905-16 and AAW70929-46 are analogous
CC to CDR regions of monoclonal antibodies specific for RSV. AAW70917-28 are
CC analogous to CDR regions of monoclonal antibodies specific for site III
CC or IV of the VP6 protein of rota virus (RV). The peptides can neutralise
CC viral infections and may also inhibit fusion between infected and
CC uninfected cells or cells and viruses. They provide passive or active
CC protection and/or inhibit transcription of the virus, so are useful as
CC antiviral agents or for prophylaxis, in human or veterinary medicine. The
CC peptides can be labelled and used to diagnose infection or contamination
CC by the virus. The peptides are particularly directed against RSV or RS
CC but may also be used against papilloma, adeno, entero, polio, influenza
CC or immune deficiency viruses.

XX Sequence 100 AA;

Query Match 100.0%; Score 46; DB 19; Length 100;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLIYWA 8

Db 39 PKLIYYA 46

RESULT 6
AAW31647
ID AAW31647 standard; Protein; 106 AA.

XX AAW31647;
XX
XX 21-MAY-1998 (first entry)
XX

DE Monoclonal antibody CP.B8 light chain variable region.

XX Cytokine receptor; gamma common chain; gc chain; human;
XX blocking agent; monoclonal antibody; CP.B8; immunological disease;
XX myasthenia gravis; rheumatoid arthritis; lupus; multiple sclerosis;
XX insulin-dependent diabetes; inflammatory bowel disease;
XX sympathetic ophthalmia; uveitis; allergy; asthma; infection;
XX graft versus host disease; psoriasis; immunosuppressive; therapy;
XX complementarity determining region; CDR.
XX
XX Mus musculus.

OS
XX
XX Key Location/Qualifiers
XX Region 24..34

FT /note= "CDR1"
FT Region 50..56
FT /note= "CDR2"
FT Region 89..97
FT /note= "CDR3"

XX WO9743416-A1.

XX 20-NOV-1997.

XX 09-MAY-1997; 97WO-US07870.

XX 10-MAY-1996; 96US-0017466.

XX (BIOJ) BIOGEN INC.

XX Benjamin CD, Burkly LC, Hession C, Whitty A;

XX WPI: 1998-008885/01.

XX N-PSDB; AAT97440.

XX Blocking agents of the gamma common chain of cytokine receptors -
XX particularly monoclonal antibodies, used to induce T cell anergy for
XX treatment of immunological diseases

XX Claim 22; Page 81; 11pp; English.

XX This polypeptide comprises the light chain variable region (VL) of
XX monoclonal antibody (Mab) C9.B8, which is produced by a hybridoma
XX deposited as ATCC 12107, and which is specific for the gamma
XX constant (gc) chain (see AAW31646) of human cytokine receptors. The
XX invention provides compositions and methods for inhibiting cytokine
XX signalling using gc chain blocking agents for the treatment of
XX immunological diseases such as myasthenia gravis, rheumatoid
XX arthritis, lupus, multiple sclerosis, insulin-dependent diabetes,
XX inflammatory bowel disease, sympathetic ophthalmia, uveitis,
XX allergy, asthma, parasitic infection, graft vs. host disease or
XX psoriasis. Preferred gc blocking agents include Mab CP.B8, its Fab
XX fragment and an antibody having a light chain variable region
XX CDR selected from those of CP.B8 VL or a heavy chain variable
XX region CDR selected from those of CP.B8 VH (see AAW31648).

XX Sequence 106 AA;

Query Match 100.0%; Score 46; DB 19; Length 106;
Best Local Similarity 100.0%; Pred. No. 0.27;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLIYYA 8
XXXXXXXXXX
Db 44 PKLIYYA 51

RESULT 7
AAR25723
ID AAR25723 standard; Protein; 107 AA.

XX AAR25723;
XX
XX 13-JAN-1993 (first entry)
XX

DE Humanised VL region of the mouse Fd 138-80 antibody.

XX Murine; immunoglobulin; CDR; HSV; non immunogenic; herpes simplex
XX virus; HSV; light chain; variable region; framework; human; Eu.
XX
XX Mus musculus.

OS
XX
XX Key Location/Qualifiers
XX Region 24..38

FT /note= "CDR"
FT Region 50..56
FT /note= "CDR"
FT Region 89..97
FT /note= "CDR"

FT Misc-difference 36 /note= "mutated residue"

FT Misc-difference 48 /note= "mutated residue"

FT Misc-difference 63 /note= "mutated residue"

FT Misc-difference 87 /note= "mutated residue"

XX WO9211018-A.

XX 09-JUL-1992.

XX 19-DEC-1991; 91WO-US09711.

XX 19-DEC-1990; 90US-0634278.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Co MS, Coellingh KL, Landolfi NF, Queen CL, Schneider WP;

XX WPI: 1992-249842/30.

XX New immunoglobulin(s) having murine CDRs in human framework
XX regions - have lower antigenicity; useful for treating e.g. HSV,
XX CMV, T-cell disorders, myeloid disorders and auto-immune
XX conditions

XX Claim 19; Fig 14D; 141pp; English.

XX The sequence shows the humanised mature light chain variable region
XX of the mouse Fd 138-80 antibody. Murine CDRs were used in a
XX human Eu framework to produce a pure humanised immunoglobulin (Ig)
XX which is capable of binding to a herpes simplex virus-specific
XX epitope. The Ig is non immunogenic, due to the human framework,
XX and has a strong affinity for its predetermined antigen. They can
XX be produced in large quantities via recombinant DNA and monoclonal
XX antibody technology. The humanised Igs may be used alone or in
XX combination with chemotherapeutic agents such as non-steroidal
XX anti-inflammatory drugs or immunosuppressants.
XX See also AAR25721-32.

XX Sequence 107 AA;

Query Match 100.0%; Score 46; DB 13; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLIYWA 8
 |||||
 Db 44 PKLIYWA 51

RESULT 8

AAR38599

ID AAR38599 standard; peptide; 107 AA.

XX AAR38599;

XX 28-OCT-1993 (first entry)

XX Human heavy chain subgroup 1 (hH1).

XX Antibody; variable domain; light; L; heavy; H; consensus;

XX affinity; antigen; immunogenicity; humanisation; framework.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 96

XX /note= "residue conserved in less than 50% of the known sequences of hH1"

XX Misc-difference 104

XX /note= "residue conserved in less than 50% of the known sequences of hH1"

XX WO9311794-A.

XX 24-JUN-1993.

XX 14-DEC-1992; 92WO-US10906.

XX 13-DEC-1991; 91US-0808464.

XX (XOMA) XOMA CORP.

XX Fishwild DM, Kohn FR, Little RG, Studnicka GW;

XX WPI; 1993-213827/26.

XX Antibodies prep. used for treatment of auto-immune diseases - by

XX replacement of critical residues to reduce immunogenicity but

XX retain binding affinity, etc.

XX Claim 2; Page 98-99; 160pp; English.

XX The consensus amino acid sequences for the subgroups of light

XX chains (hK1 - AAR38590, hK3 - NGK, hK2 - GST, hL1 - AAR38591, hL2 -

XX AAR38592, hL3 - AAR38593, hL6 - AAR38594, hK4 - AAR38595, hL4 -

XX AAR38596, and hL5 - AAR38597) and heavy chains (hH3 - AAR38598, hH1 -

XX AAR38599 and hH2 - AAR38600) of human variable domains may be used to

XX prepare, for example, a modified mouse antibody variable domain that

XX retains the affinity of the natural domain for antigen while exhibiting

XX reduced immunogenicity in humans.

XX Unlike other methods of humanisation, which advocate the

XX replacement of entire antibody framework regions with those of human

XX antibodies, this method involves only the introduction of human

XX residues into those positions not critical for antigen binding.

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PKLIYWA 8
 |||||
 Db 45 PKLIYWA 52

RESULT 9

AAW80817

ID AAW80817 standard; Protein; 107 AA.

XX AAW80817;

XX 16-FEB-1999 (first entry)

XX Amino acid sequence of human kappa 5.1 light chain variable region.

XX Human; kappa 5.1 light chain variable region; receptor; antigen;

XX tumour; auto-immune disease; graft rejection; allergy;

XX inflammatory disease; endocrine disease; degenerative disease.

XX Homo sapiens.

XX WO9846645-A2.

XX 22-OCT-1998.

XX 14-APR-1998; 98WO-EP02180.

XX 14-APR-1997; 97EP-0106109.

XX (KUFE/) KUFE P.

XX (RAUM/) RAUM T.

XX Kufer P, Raum T;

XX WPI; 1998-594564/50.

XX N-PSDB; AAW68539.

XX Production of anti-human antigen receptors - by selecting a

XX combination of functionally rearranged VH and VL immunoglobulin

XX chains expressed from a recombinant vector

XX Claim 9; Fig 9; 84pp; English.

XX This is the amino acid sequence of the human kappa light chain

XX variable region, used in the method of the invention, for providing

XX receptors that can be used for targeting antigens in humans without

XX being immunogenic themselves. Such receptors can be used for treating

XX diseases such as tumours or auto-immune diseases, graft rejection

XX after transplantation, infectious diseases by targeting cellular

XX receptors as well as allergic, inflammatory, endocrine and

XX degenerative diseases by targeting key molecules involved in the

XX pathological processes.

XX Sequence 107 AA;

Query Match 100.0%; Score 46; DB 19; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10

AAW80814

ID AAW80814 standard; Protein; 107 AA.

XX AAW80814;

DT 16-FEB-1999 (first entry)
 XX Amino acid sequence of human Kappa 8 light chain variable region.
 DE
 XX
 XX
 KM Human: kappa 8 light chain variable region; receptor; antigen;
 KM tumour; auto-immune disease; graft rejection; allergy;
 KM Inflammatory disease; endocrine disease; degenerative disease.
 XX
 OS Homo sapiens.
 XX
 PN WO9846645-A2.
 XX
 PD 22-OCT-1998.
 XX
 XX 14-APR-1998; 98WO-EP02180.
 PF
 XX 14-APR-1997; 97EP-0106109.
 PR
 XX (KUPE/) KUPEP P.
 PA (RAUM/) RAUM T.
 XX
 PI Kufer P, Raum T;
 XX
 DR WPI: 1998-594564/50.
 DR N-PSDB; AAV68536.
 XX
 PT Production of anti-human antigen receptors - by selecting a
 PT combination of functionally rearranged VH and VL immunoglobulin
 PT chains expressed from a recombinant vector
 XX
 PS Claim 9; Fig 6; 84pp; English.
 XX
 CC This is the protein sequence of the human kappa 8 light chain
 CC variable region, used in the method of the invention, for providing
 CC receptors that can be used for targeting antigens in humans without
 CC being immunogenic themselves. Such receptors can be used for treating
 CC diseases such as tumours or auto-immune diseases, graft rejection
 CC after transplantation, infectious diseases by targeting cellular
 CC receptors as well as allergic, inflammatory, endocrine and
 CC degenerative diseases by targeting key molecules involved in the
 CC pathological process.
 CC
 SQ Sequence 107 AA;
 XX

Query Match 100.0%; Score 46; DB 19; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRLIYWA 8
 |||||
 Db 44 PRLIYWA 51

RESULT 11
 AAW58499
 ID AAW58499 standard; protein; 107 AA.
 XX
 AC AAW58499;
 XX
 DT 18-AUG-1998 (first entry)
 XX
 DE Human kappa light chain subgroup 4 consensus sequence hK4.
 XX
 KW Humanised; human; mouse; CD5; anti-CD5 antibody; immunoglobulin;
 KW depletion; cytotoxic; immunconjugate; fusion protein; psoriasis;
 KW autoimmune disease; rheumatoid arthritis; type I diabetes.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 96

FT
 FT Misc-difference 104 /note= "not specified"
 FT /note= "not specified"
 XX
 XX US5770196-A.
 PN
 PD 23-JUN-1998.
 XX
 XX 07-JUN-1995; 95US-0472788.
 PF
 XX 23-JUN-1993; 93US-0082842.
 PR 13-DEC-1991; 91US-0808464.
 PR 14-DEC-1992; 92WO-US10906.
 PR 07-JUN-1995; 95US-0472788.
 XX
 PA (XOMA) XOMA CORP.
 XX
 XX Studnicka GM;
 PI
 DR WPI: 1998-376744/32.
 XX
 PT Depletion of CD5-positive cells in vivo - using anti-CD5 antibodies
 PT with humanised variable regions
 XX
 PS Disclosure; Column 57-58; 77pp; English.
 XX
 CC A method has been developed of depleting CD5+ cells in an animal. The
 CC method comprises administering a cytotoxic protein containing a modified
 CC immunoglobulin (Ig) variable domain, where the protein is an anti-CD5
 CC molecule or an immunconjugate or fusion protein containing an anti-CD5
 CC Ig molecule, and where the modified Ig variable domain comprises at
 CC least one of (a) a modified light chain variable region (see AAW58478 or
 CC AAW58480), and (b) a modified heavy chain variable region (see AAW58479
 CC or AAW58481), where AAW58478 and AAW58479 are humanised forms of the H65
 CC light and heavy chain variable domains with low risk amino acid
 CC substitutions [i.e. low risk of reducing antigen-binding specificity.]
 CC and AAW58480 and AAW58481 are humanised forms of the H65 light and heavy
 CC chain variable domains with moderate risk amino acid substitutions and
 CC are present in humanised H65 antibody h63 (ATCC HB 11206). The method is
 CC useful for treating autoimmune diseases, especially systemic lupus
 CC erythematosus, rheumatoid arthritis, psoriasis or type I diabetes. The
 CC present sequence represents a consensus amino acid sequence for light
 CC chain subgroups of human antibody variable domains, from the present
 CC invention.
 CC
 SQ Sequence 107 AA;
 XX

Query Match 100.0%; Score 46; DB 19; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PRLIYWA 8
 |||||
 Db 45 PRLIYWA 52

RESULT 12
 AAW47087
 ID AAW47087 standard; protein; 107 AA.
 XX
 AC AAW47087;
 XX
 DT 26-JUN-1998 (first entry)
 XX
 DE Mouse J591 monoclonal antibody light chain variable region VK17.
 DE Mouse J591 monoclonal antibody light chain variable region VK17.
 XX
 KW Mouse; monoclonal antibody; J591; prostate specific membrane antigen;
 KW cancer; vascular endothelial cell; metastatic adenocarcinoma.
 XX
 OS Mus sp.
 OS
 XX WO9803873-A1.
 PN

XX 29-JAN-1998.
 XX PD
 XX 17-JUL-1997; 97WO-US12035.
 XX PF
 XX 09-APR-1997; 97US-0838682.
 XX PR 18-JUL-1996; 96US-0022125.
 XX PA (CORR) CORNELL RES FOUND INC.
 XX PI Bander NH;
 XX WPI: 1998-120937/11.
 XX DR N-PSDB; AAV13953.
 XX PT Destroying cancer cells with agent that binds to prostate specific
 XX PT membrane antigen - on vascular endothelial cells near the cancer, or
 XX PT on normal, hypertrophic or cancerous prostatic cells, also used for
 XX PT diagnosis
 XX PS Example 12; Page 61; 94pp; English.
 XX CC The present sequence represents the mouse J591 monoclonal antibody light
 XX CC chain variable region from clone VK17 from an example of the present
 XX CC invention. The present invention describes the elimination of cancer
 XX CC cells by treating vascular endothelial cells (VEC) close to the cancer
 XX CC with an agent (A) able to bind to the extracellular domain (ECD) of
 XX CC prostate specific membrane antigen (PSMA). (A) both binds to the VEC and
 XX CC destroys the cancer cells. Also described are: (1) the detection of
 XX CC cancer tissue by detecting binding of labelled (A) to VEC close to, or
 XX CC within, a cancer tissue; (2) eliminating or detecting normal, benignly
 XX CC hyperplastic or cancerous prostate epithelial cells using optionally
 XX CC labelled (A); (3) hybridomas that produce a monoclonal antibody (Mab)
 XX CC that binds to PSMA. The method is used to treat renal, urothelial,
 XX CC colon, lung, rectal or breast cancers and metastatic adenocarcinoma of
 XX CC the liver. The diagnostic method is particularly used to detect
 XX CC recurrence of prostatic disease or to monitor the effect of treatments
 XX CC for prostate cancer (presence of PSMA in the serum indicates that
 XX CC prostate cells are being lysed). (A) binds to an epitope of PSMA
 XX CC expressed on live cells (contrast antibody 7E11 which only binds after
 XX CC cell lysis), allowing targeting of live, unfixed cells and thus
 XX CC providing more efficient treatment and diagnosis. Both cancer cells
 XX CC themselves and the VEC on which they depend are killed. All VEC close to
 XX CC cancer cells express PSMA, whatever the type of cancer, but normal VEC
 XX CC do not.
 XX SQ Sequence 107 AA;
 XX QY 1 PKLLIYMA 8
 XX DB 44 PKLLIYMA 51
 XX
 XX RESULT 13
 XX ID AAY17955 standard; Protein: 107 AA.
 XX AC AAY17955;
 XX XX
 XX DT 04-AUG-1999 (first entry)
 XX XX
 XX DE Human kappa 8 light chain variable region.
 XX XX
 XX KM Binding site domain; BSD; epitope; fusion protein; therapeutic; cancer;
 XX KM autoimmune disease; scFv-antibody; single-chain Fv.
 XX XX
 XX OS Homo sapiens.
 XX XX

PN WO925818-A1.
 XX XX
 XX PD 27-MAY-1999.
 XX XX
 XX PF 16-NOV-1998; 98WO-EP07313.
 XX XX
 XX PR 17-NOV-1997; 97EP-0120096.
 XX XX
 XX PA (KUFE/) KUFER P.
 XX XX
 XX PI Borschert K, Kufer P, Lutterbuese R, Raum T, Zettl F;
 XX XX
 XX DR WPI: 1999-338004/28.
 XX DR N-PSDB; AAX77237.
 XX XX
 XX PT Phase display system for identification of binding site domains
 XX PT retaining capacity to bind an epitope
 XX PT
 XX PS Disclosure; Fig 3.2; 152pp; English.
 XX XX
 XX CC The invention relates to a method of identifying binding site domains
 XX CC (BSD) that retain the capacity of binding to a predetermined epitope when
 XX CC positioned C-terminal of at least one further domain in a recombinant bi-
 XX CC or multivalent polypeptide. The method comprises (a) testing a panel of
 XX CC BSD displayed on the surface of a biological display system as part of a
 XX CC fusion protein for binding to a predetermined epitope, where the fusion
 XX CC protein comprises an additional domain positioned N-terminal of the BSD
 XX CC and an amino acid sequence that mediates anchoring of the fusion protein
 XX CC to the surface of the display system; and (b) identifying a BSD that
 XX CC binds to the predetermined epitope. The method is useful to identify bi-
 XX CC or multivalent polypeptides that comprise antibody binding sites capable
 XX CC of efficiently binding to the corresponding antigen. The polypeptides or
 XX CC antibodies identified by the method are useful therapeutically and
 XX CC diagnostically, for e.g. cancer and autoimmune diseases. scFv-antibody
 XX CC fragments that bind independently of their position within bifunctional
 XX CC single-chain fusion proteins can be isolated from combinatorial antibody
 XX CC libraries using the new in vitro method.
 XX SQ Sequence 107 AA;
 XX QY 1 PKLLIYMA 8
 XX DB 44 PKLLIYMA 51
 XX
 XX RESULT 14
 XX ID AAY90375 standard; Protein: 107 AA.
 XX AC AAY90375;
 XX XX
 XX DT 15-JAN-2001 (first entry)
 XX XX
 XX DE J591 monoclonal antibody light chain protein sequence fragment.
 XX XX
 XX KM J591 monoclonal antibody; extracellular domain; diagnosis; therapy;
 XX KM prostate specific membrane antigen; prostate cancer; light chain.
 XX XX
 XX OS Homo sapiens.
 XX XX
 XX PN US6107090-A.
 XX PN
 XX PD 22-AUG-2000.
 XX XX
 XX PF 09-APR-1997; 97US-0838682.
 XX PF
 XX PR 06-MAY-1996; 96US-0016976.
 XX PR 18-JUL-1996; 96US-0022125.
 XX XX

XX (CORR) CORNELL RES FOUND INC.
 XX Bander NH;
 PI
 XX WPI: 2000-571325/53.
 DR N-PSDB; AAA37835.
 XX
 PT Antibody specific for extracellular prostate-specific membrane antigen,
 PT useful for diagnosis and treatment of prostate cancer
 XX
 PS Example 12; Column 24; 33pp: English.
 XX
 CC This sequence is a fragment of the light chain of the monoclonal antibody
 CC J591. The invention relates to an isolated antibody or its antigen
 CC binding portion (1) which binds to an extracellular domain of prostate
 CC specific membrane antigen and which does not require cell lysis to bind
 CC to the extracellular domain. The antibody or its antigen binding portion
 CC is selected for its ability to bind to live cells. (1) is useful for
 CC diagnosis of diseases associated with the presence of normal, benign
 CC hyperplastic, and cancerous epithelial cells or portions. Also it can be
 CC used for identifying the recurrence of such diseases, particularly when
 CC the disease is localised in a particular biological material of the
 CC patient for e.g. recurrence of prostatic disease. They can also be used
 CC alone or bound to a substance effective to kill cancerous prostate
 CC epithelial cells as a therapy for prostate cancer. Binding and
 CC internalising of the antibody with the prostate specific membrane
 CC antigen, permits the therapeutic use of intracellularly acting cytotoxic
 CC agents. (1) targets only prostate epithelial cells and other tissue are
 CC spared which provides safer treatment particularly for elderly patients.
 CC The antibodies bind to living prostate cells and treatments using these
 CC antibodies are more effective than those which target lysed prostate
 CC cells.
 XX
 SQ Sequence 107 AA;
 XX
 Query Match 100.0%; Score 46; DB 21; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PKLLIYWA 8
 DB 44 PKLLIYWA 51
 XX
 RESULT 15
 ID AAY79425
 XX AAY79425 standard: Protein; 107 AA.
 AC AAY79425;
 XX
 DT 01-AUG-2000 (first entry)
 XX
 DE Tie2 receptor antagonist antibody 12H8 light chain variable region.
 XX
 KW Tie2 receptor; tyrosine kinase receptor; antagonist;
 KW monoclonal antibody; 12H8; light chain; mouse; angiogenesis;
 KW cancer; diabetic retinopathy; macular degeneration; arthritis;
 KW psoriasis; atherosclerosis; bone marrow; gene therapy; inhibitor;
 KW antiproliferative; antiinflammatory; proliferative;
 KW antithrombotic; antianemic.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Region 24..34
 FT /note= "complementarity determining region 1"
 FT Region 50..56
 FT /note= "complementarity determining region 2"
 FT Region 89..97
 FT /note= "complementarity determining region 3"
 XX

PN WO200018437-A1.
 XX
 PD 06-APR-2000.
 XX
 PF 28-SEP-1999; 99WO-US22421.
 XX
 PR 28-SEP-1998; 98US-0102100.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Holmes SD, Erickson-Miller CL, Winkler JD;
 XX
 DR WPI: 2000-303392/26.
 DR N-PSDB; AA294887.
 XX
 PT Novel Tie receptor antagonist antibodies inhibits angiogenesis useful
 PT for treating e.g. diabetic retinopathy, cancer or for expanding blood
 PT cells for transplant
 XX
 PS Claim 10; Page 35; 42pp: English.
 XX
 CC The present sequence is that of the light chain variable region of
 CC murine monoclonal antibody 12H8, as deduced from isolated cDNA
 CC clones (see AA294887). 12H8 is an antagonist antibody of the Tie2
 CC receptor; the antibody binds to the Tie2 receptor and inhibits
 CC angiogenesis. The Tie2 receptor antagonist antibody (which may be
 CC humanized) is useful for inhibiting angiogenesis in a mammal for:
 CC (1) treating an angiogenic disease especially diabetic retinopathy
 CC or macular degeneration; (2) treating proliferative diseases such as
 CC cancer, arthritis, psoriasis or atherosclerosis; (3) expansion of
 CC marrow cells transduced with a gene for gene therapy; (4)
 CC expanding cord blood, marrow or peripheral blood cell for
 CC transplant; (5) enhancing the survival of hematopoietic progenitor
 CC cells; and (6) to increase the proliferation of megakaryotic
 CC (progenitor) cells for increasing platelet numbers (all claimed).
 XX
 SQ Sequence 107 AA;
 XX
 Query Match 100.0%; Score 46; DB 21; Length 107;
 Best Local Similarity 100.0%; Pred. No. 0.27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PKLLIYWA 8
 DB 44 PKLLIYWA 51
 XX

Search completed: July 15, 2002, 12:57:56
 Job time: 412 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 12:59:17 ; Search time 75.67 Seconds
(without alignments)
2.582 Million cell updates/sec

Title: US-09-712-819A-4

Perfect score: 46

Sequence: 1 PKLIYNA 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6CTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	100.0	32	3	US-08-513-968-55	Sequence 55, Appl
2	46	100.0	107	1	US-07-634-278-50	Sequence 50, Appl
3	46	100.0	107	1	US-07-634-278-51	Sequence 51, Appl
4	46	100.0	107	1	US-08-477-728-50	Sequence 50, Appl
5	46	100.0	107	1	US-08-477-728-51	Sequence 51, Appl
6	46	100.0	107	1	US-08-474-040-50	Sequence 50, Appl
7	46	100.0	107	1	US-08-474-040-51	Sequence 51, Appl
8	46	100.0	107	1	US-08-487-200-50	Sequence 50, Appl
9	46	100.0	107	1	US-08-487-200-51	Sequence 51, Appl
10	46	100.0	107	1	US-08-488-113B-156	Sequence 156, App
11	46	100.0	107	1	US-08-477-484B-156	Sequence 156, App
12	46	100.0	107	1	US-08-107-669D-20	Sequence 20, Appl
13	46	100.0	107	1	US-08-472-788A-20	Sequence 20, Appl
14	46	100.0	107	2	US-08-477-531B-20	Sequence 20, Appl
15	46	100.0	107	2	US-08-646-360-156	Sequence 156, App
16	46	100.0	107	2	US-08-082-842A-20	Sequence 20, Appl
17	46	100.0	107	3	US-08-838-682-19	Sequence 19, Appl
18	46	100.0	107	4	US-08-895-914-19	Sequence 19, Appl
19	46	100.0	107	4	US-08-839-765-156	Sequence 156, App
20	46	100.0	107	4	US-09-136-389-156	Sequence 156, App
21	46	100.0	107	4	US-08-484-537-50	Sequence 50, Appl
22	46	100.0	107	4	US-08-484-537-51	Sequence 51, Appl
23	46	100.0	107	4	US-09-357-710A-19	Sequence 19, Appl
24	46	100.0	109	2	US-08-308-494A-23	Sequence 23, Appl
25	46	100.0	110	4	US-08-957-001B-5	Sequence 5, Appl1
26	46	100.0	110	4	US-08-957-001B-24	Sequence 24, Appl1
27	46	100.0	110	4	US-09-496-301-5	Sequence 5, Appl1

28	46	100.0	110	4	US-09-496-301-24	Sequence 24, Appl
29	46	100.0	112	2	US-07-942-245-30	Sequence 30, Appl
30	46	100.0	112	1	US-07-916-098A-15	Sequence 15, Appl
31	46	100.0	113	1	US-08-690-102A-2	Sequence 2, Appl1
32	46	100.0	113	1	US-08-690-102A-6	Sequence 6, Appl1
33	46	100.0	113	2	US-08-263-911-2	Sequence 2, Appl1
34	46	100.0	113	3	US-08-819-033-1	Sequence 1, Appl1
35	46	100.0	113	3	US-08-483-749A-16	Sequence 16, Appl
36	46	100.0	113	3	US-08-513-968-82	Sequence 82, Appl
37	46	100.0	113	4	US-09-127-902-2	Sequence 2, Appl1
38	46	100.0	113	4	US-09-127-902-6	Sequence 6, Appl1
39	46	100.0	113	4	US-09-153-107-2	Sequence 2, Appl1
40	46	100.0	113	4	US-09-153-107-6	Sequence 6, Appl1
41	46	100.0	113	4	US-09-155-107-20	Sequence 20, Appl
42	46	100.0	113	4	US-08-525-539A-80	Sequence 80, Appl
43	46	100.0	113	4	US-09-025-203-7	Sequence 7, Appl1
44	46	100.0	113	4	US-09-025-203-8	Sequence 8, Appl1
45	46	100.0	113	4	US-09-025-203-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-08-513-968-55
; Sequence 55, Application US/08513968
; Patent No. 6114143
; GENERAL INFORMATION:
; APPLICANT: EDA, Yasuyuki
; APPLICANT: MAEDA, Hiroaki
; APPLICANT: MAKIZUMI, Keiichi
; APPLICANT: SHIOSAKI, Keiichi
; APPLICANT: OSATOMI, Kiyoshi
; APPLICANT: KIMACHI, Kazuhiko
; APPLICANT: HIGUCHI, Hirofumi
; APPLICANT: TOKIYOSHI, Sachio
; TITLE OF INVENTION: ANTI-HIV MONOCLONAL ANTIBODY
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,968
; FILING DATE: 11-SEP-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 78913/1993
; FILING DATE: 11-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: EDA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-513-968-55

Query Match 100.0%; Score 46; DB 3; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
DB 8 PKLLIYWA 15

RESULT 2
US-07-634-278-50
; Sequence 50, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas P.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-634-278-50

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
DB 44 PKLLIYWA 51

RESULT 3
US-07-634-278-51
; Sequence 51, Application US/07634278
; Patent No. 5530101
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas P.
; APPLICANT: COELINGH, Kathleen L.
; APPLICANT: SELICK, Harold E.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/634,278
; FILING DATE: 19-DEC-1990
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/590,274
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/310,252
; FILING DATE: 13-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/290,975
; FILING DATE: 28-DEC-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 11823-002600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-634-278-51

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
DB 44 PKLLIYWA 51

RESULT 4
US-08-477-728-50
; Sequence 50, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: SELICK, Harold E.

;; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
;; NUMBER OF SEQUENCES: 113
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Townsend and Townsend and Crew LLP
;; STREET: Two Embarcadero Center, 8th floor
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: US
;; ZIP: 94111
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,728
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/310,252
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/290,975
;; FILING DATE: 28-DEC-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 11823-002600
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 50:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 107 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-477-728-50

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
DB 44 PKLLIYWA 51

RESULT 5
US-08-477-728-51
; Sequence 51, Application US/08477728
; Patent No. 5585089
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: SCHNEIDER, William P.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th floor
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94111

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/477,728
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/634,278
;; FILING DATE: 19-DEC-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/590,274
;; FILING DATE: 28-SEP-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/310,252
;; FILING DATE: 13-FEB-1989
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/290,975
;; FILING DATE: 28-DEC-1988
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Smith, William M
;; REGISTRATION NUMBER: 30,223
;; REFERENCE/DOCKET NUMBER: 11823-002600
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 326-2400
;; TELEFAX: (415) 326-2422
;; INFORMATION FOR SEQ ID NO: 51:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 107 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-477-728-51

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYWA 8
DB 44 PKLLIYWA 51

RESULT 6
US-08-474-040-50
; Sequence 50, Application US/08474040
; Patent No. 5693761
; GENERAL INFORMATION:
; APPLICANT: QUEEN, Cary L.
; APPLICANT: CO, Man Sung
; APPLICANT: SCHNEIDER, William P.
; APPLICANT: LANDOLFI, Nicholas F.
; APPLICANT: GOELINGH, Kathleen L.
; TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,040
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-040-50

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLIYWA 8
DB 44 PKLIYWA 51

RESULT 7
US-08-474-040-51
Sequence 51, Application US/08474040
Patent No. 5693761
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,040
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274

FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002600
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-040-51

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PKLIYWA 8
DB 44 PKLIYWA 51

RESULT 8
US-08-487-200-50
Sequence 50, Application US/08487200
Patent No. 5693762
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: CO, Man Sung
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975

FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30, 223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-487-200-50

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
DB 44 PKLLIYMA 51

RESULT 9
US-08-487-200-51
Sequence 51, Application US/08487200
Patent No. 5693762
GENERAL INFORMATION:
APPLICANT: QUEEN, Cary L.
APPLICANT: SCHNEIDER, William P.
APPLICANT: LANDOLFI, Nicholas F.
APPLICANT: COELINGH, Kathleen L.
APPLICANT: SELICK, Harold E.
TITLE OF INVENTION: IMPROVED HUMANIZED IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,200
FILING DATE: 7-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/634,278
FILING DATE: 19-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/590,274
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/310,252
FILING DATE: 13-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/290,975
FILING DATE: 28-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 11823-002610
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-487-200-51

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
DB 44 PKLLIYMA 51

RESULT 10
US-08-488-113B-156
Sequence 156, Application US/08488113B
Patent No. 5744580
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Studlika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
TITLE OF INVENTION: Proteins
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,113B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 110220S07/200-70.P3.C2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-9889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-113B-156

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
|||||||
DB 45 PKLIYWA 52

RESULT 11

US-08-477-484B-156
Sequence 156, Application US/08477484B
Patent No. 5756699
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studlika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477.484B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-9155
TELEFAX: 312/707-8889
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-477-484B-156

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
|||||||
DB 45 PKLIYWA 52

RESULT 12

US-08-107-669D-20
Sequence 20, Application US/08107669D
Patent No. 5766886
GENERAL INFORMATION:
APPLICANT: Studlika, Gary M.
TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/107.669D
FILING DATE: 13-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10906
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,464
FILING DATE: 13-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Michele A. Cimbalia
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0610.1000001/MAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/371-2600
TELEFAX: 202/371-2540
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-107-669D-20

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLIYWA 8
|||||||
DB 45 PKLIYWA 52

RESULT 13
US-08-472-788A-20
Sequence 20, Application US/08472788A
Patent No. 5770196
GENERAL INFORMATION:

APPLICANT: Studnicka, Gary M.
TITLE OF INVENTION: Modified Antibody Variable Domains
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,788A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/082,842
FILING DATE: 23-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10906
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,464
FILING DATE: 13-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0610.1000003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/371-2600
TELEFAX: 202/371-2540
TELEX:
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-788A-20

Query Match 100.0%; Score 46; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
|||||
DB 45 PKLLIYMA 52

RESULT 14
US-08-477-531B-20
Sequence 20, Application US/08477531B
Patent No. 5821123
GENERAL INFORMATION:
APPLICANT: Studnicka, Gary M.
TITLE OF INVENTION: Modified Antibody Variable Domains (as amended)
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein and Fox P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,531B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/107,669
FILING DATE: 13-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10906
FILING DATE: 14-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/808,464
FILING DATE: 13-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Michele A. Cimbala
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0610.1000004/MAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/371-2600
TELEFAX: 202/371-2540
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-531B-20

Query Match 100.0%; Score 46; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PKLLIYMA 8
|||||
DB 45 PKLLIYMA 52

RESULT 15
US-08-646-360-156
Sequence 156, Application US/08646360
Patent No. 5837491
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroli, Stephen F.
APPLICANT: Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: MCNICHOLAS, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 156:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-646-360-156

Query Match 100.0%; Score 46; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PKLLIYWA 8
|||||||
DB 45 PKLLIYWA 52

Search completed: July 15, 2002, 12:59:18
Job time: 389 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:01:01 ; Search time 95.45 Seconds
(without alignments)
7.047 Million cell updates/sec

Title: US-09-712-819A-6

Perfect score: 33

Sequence: 1 FTLKISR 7

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	54	2 S34093	Ig kappa chain V r
2	33	100.0	71	2 H30538	Ig kappa chain V r
3	33	100.0	72	2 S40358	Ig kappa chain - h
4	33	100.0	75	2 S40337	Ig kappa chain V-J
5	33	100.0	83	2 S34095	Ig kappa chain V r
6	33	100.0	83	2 S24211	Ig kappa chain V r
7	33	100.0	87	2 S34091	Ig kappa chain V r
8	33	100.0	87	2 S34094	Ig kappa chain V r
9	33	100.0	89	2 B25155	Ig kappa chain V r
10	33	100.0	91	2 S42186	Ig kappa chain V r
11	33	100.0	91	2 S25463	Ig kappa chain V r
12	33	100.0	93	2 PH1039	Ig light chain V r
13	33	100.0	94	2 PH0258	Ig kappa chain V r
14	33	100.0	96	2 S40320	Ig kappa chain - h
15	33	100.0	99	2 S24501	Ig kappa chain V r
16	33	100.0	99	2 E28833	Ig heavy chain V r
17	33	100.0	99	2 S24504	Ig kappa chain V r
18	33	100.0	100	2 S26334	Ig light chain V r
19	33	100.0	100	2 S24681	Ig kappa chain - h
20	33	100.0	100	2 S45440	Ig kappa chain V r
21	33	100.0	101	2 PH1057	Ig kappa chain V r
22	33	100.0	101	2 C34153	Ig kappa chain V-I
23	33	100.0	101	2 A33730	Ig kappa chain V r
24	33	100.0	102	2 PH1035	Ig light chain V r
25	33	100.0	102	2 PH1044	Ig light chain V r
26	33	100.0	102	2 S14592	Ig kappa chain V r
27	33	100.0	102	2 E28195	Ig kappa chain V r
28	33	100.0	102	2 S14594	Ig kappa chain V r
29	33	100.0	102	2 S14593	Ig kappa chain V r

30	33	100.0	102	2 S14591	Ig kappa chain V r
31	33	100.0	102	2 S14590	Ig kappa chain V r
32	33	100.0	103	2 PH1034	Ig light chain V r
33	33	100.0	103	2 PH1036	Ig light chain V r
34	33	100.0	103	2 PH1037	Ig light chain V r
35	33	100.0	103	2 PH1038	Ig light chain V r
36	33	100.0	103	2 PH1040	Ig light chain V r
37	33	100.0	103	2 PH1041	Ig light chain V r
38	33	100.0	103	2 PH1042	Ig light chain V r
39	33	100.0	103	2 PH1043	Ig light chain V r
40	33	100.0	103	2 PH1045	Ig light chain V r
41	33	100.0	103	2 PH1055	Ig light chain V r
42	33	100.0	103	2 PH1056	Ig light chain V r
43	33	100.0	103	2 PH1098	Ig light chain V r
44	33	100.0	103	2 PH1099	Ig light chain V r
45	33	100.0	103	2 PH1100	Ig light chain V r

ALIGNMENTS

RESULT 1
S34093
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 07-May-1999
C:Accession: S34093
R:Wagner, S.D.; Iuzza, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distrib
A:Reference number: S34076; MUID:93170387
A:Accession: S34093
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-54 <MAC>
A:Cross-references: EMBL:X67177
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1992
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
DB 43 FTLKISR 49

RESULT 2
H30538
Ig kappa chain V region (174-3f4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 10-Feb-1989 #sequence_revision 10-Feb-1989 #text_change 09-May-1997
C:Accession: H30538
R:Claflin, J.L.; Berry, J.
J. Immunol. 141, 4012-4019, 1988
A:Title: Genetics of the phosphocholine-specific antibody response to Streptococcus p
A:Reference number: A30534; MUID:89035545
A:Accession: H30538
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-71 <CLAS>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7

```
DB      35  FTLKISR 41
|||||
RESULT  3
S40358
Ig kappa chain - human
C:Species: Homo sapiens (man)
C:Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S40358
R:Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chl genes and their hypermutation.
A:Reference number: S40312; MUID:94080891
A:Accession: S40358
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-72 <KLE>
A:Cross-references: EMBL:X72468; NID:q441404; PIDN:CA451136.1; PID:q441405
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  FTLKISR 7
|||||
DB      28  FTLKISR 34

RESULT  4
S40337
Ig kappa chain V-J-C region - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S40337
R:Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chl genes and their hypermutation.
A:Reference number: S40312; MUID:94080891
A:Accession: S40337
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-75 <KLE>
A:Cross-references: EMBL:X72447; NID:q441362; PIDN:CA451115.1; PID:q441363
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  FTLKISR 7
|||||
DB      30  FTLKISR 36

RESULT  5
S34095
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34095
R:Magner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34095
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-83 <WAG>
A:Cross-references: EMBL:X67179
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
F:1-80/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  FTLKISR 7
|||||
DB      61  FTLKISR 67

RESULT  6
S24211
Ig kappa chain V region (V3a) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
C:Accession: S24211
R:Pargent, W.; Meindl, A.; Thiebe, R.; Mitzel, S.; Zachau, H.G.
Eur. J. Immunol. 21, 1821-1827, 1991
A:Title: The human immunoglobulin kappa locus. Characterization of the duplicated O r
A:Reference number: S24205; MUID:93330953
A:Accession: S24211
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <PAR>
A:Cross-references: EMBL:X59317; NID:q33270; PIDN:CA442004.1; PID:q33271
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1  FTLKISR 7
|||||
DB      59  FTLKISR 65

RESULT  7
S34091
Ig kappa chain V region (patient 19 and 20) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Mar-1998 #text_change 21-Jan-2000
C:Accession: S34091; S34092
R:Magner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distrib
A:Reference number: S34076; MUID:93170387
A:Accession: S34091
A:Molecule type: DNA
A:Residues: 1-87 <WAG>
A:Cross-references: EMBL:X67175
A:Experimental source: patient 19
A:Accession: S34092
A:Molecule type: DNA
A:Residues: 1-87 <WAG>
A:Cross-references: EMBL:X67176
A:Experimental source: patient 20
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:3-82/Domain: Immunoglobulin homology <IMM>

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 87;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 FTLKISR 7
|||||||
Db 63 FTLKISR 69

RESULT 8
S34094
Ig kappa chain V region - human
C:Species: Homo sapiens (man)
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S34094
R:Wagner, S.D.; Luzzatto, L.
Eur. J. Immunol. 23, 391-397, 1993
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed
A:Reference number: S34076; MUID:93170387
A:Accession: S34094
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <NAG>
A:Cross-references: EMBL:X67178
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:3-82/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 33; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 63 FTLKISR 69

RESULT 9
B25155
Ig kappa chain V region (SML.5) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 28-Dec-1987 #sequence_revision 02-Jun-1988 #text_change 16-Aug-1996
C:Accession: B25155
R:Robbins, P.F.; Rosen, E.M.; Haba, S.; Nisonoff, A.
Proc. Natl. Acad. Sci. U.S.A. 83, 1050-1054, 1986
A:Title: Relationship of V(H) and V(L) genes encoding three idiotypic families of anti-F
A:Reference number: A94083; MUID:86149212
A:Accession: B25155
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-89 <ROB>
C:Comment: This chain is expressed in an IGM with anti-arsenate activity.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 54 FTLKISR 60

RESULT 10
S42186
Ig kappa chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 28-Sep-1994 #sequence_revision 21-Jul-1995 #text_change 21-Jan-2000
C:Accession: S42186; S42195
R:Mo, J.A.; Bona, C.A.; Holmdahl, R.
Eur. J. Immunol. 23, 2503-2510, 1993
A:Title: Variable region gene selection of immunoglobulin G-expressing B cells with spec
A:Reference number: S42176; MUID:94009207
A:Accession: S42186

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <MOJ>
A:Cross-references: EMBL:Z25442; NID:9407830; PIDN:CAA80929.1; PID:9407831
A>Note: The authors translated the codon GTT for residue 81 as Thr, TGG for residue 8
A:Accession: S42195
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <MO2>
A:Cross-references: EMBL:Z25460; NID:9407848; PIDN:CAA80947.1; PID:9407849
A>Note: The authors translated the codon GTT for residue 81 as Thr, TGG for residue 8
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:7-86/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 33; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 67 FTLKISR 73

RESULT 11
S25463
Ig kappa chain V region - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Jul-1999
C:Accession: S25463
R:Nonestier, M.; Fasy, T.M.; Losman, M.J.; Novick, K.E.; Muller, S.
submitted to the EMBL Data Library July 1992
A:Description: Structure and binding properties of monoclonal antibodies to core hist
A:Reference number: S25174
A:Accession: S25463
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-91 <MON>
A:Cross-references: EMBL:X67624; NID:952189; PIDN:CAA47882.1; PID:9338264
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 76 FTLKISR 82

RESULT 12
PH1039
Ig light chain V region (clone 165.27) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Jun-1996
C:Accession: PH1039
R:Tillman, D.M.; Jou, N.T.; Hill, R.J.; Marlon, T.N.
J. Exp. Med. 176, 761-779, 1992
A:Title: Both Igm and IgG anti-DNA antibodies are the products of clonally selective
A:Reference number: PH0971; MUID:92381444
A:Accession: PH1039
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-93 <TIL>
A:Experimental source: B cell, strain [NZB x NZW]F1
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 93;

Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 67 FTLKISR 73

RESULT 13

PI0258
Ig kappa chain V region (anti-DNA, 1E81VK) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Aug-1996
C:Accession: PI0258
R:Shlomchik, M.; Masceilli, M.; Shan, H.; Radic, M.Z.; Pisetsky, D.; Marshak-Rochestein, A.
J. Exp. Med. 171, 265-297, 1990
A:Title: Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic
A:Reference number: PI0231; MUID:90111618
A:Accession: PI0258
A:Molecule type: mRNA
A:Residues: 1-94 <SHL>
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin
F:1-6/Region: framework 1
F:7-22/Region: complementarity-determining 1
F:23-37/Region: framework 2
F:38-44/Region: complementarity-determining 2
F:45-76/Region: framework 3
F:77-85/Region: complementarity-determining 3
F:86-94/Region: framework 4

Query Match 100.0%; Score 33; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 59 FTLKISR 65

RESULT 14

S40320
Ig kappa chain - human
C:Species: Homo sapiens (man)
C:Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 23-Jul-1999
C:Accession: S40320
R:Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A:Title: Expressed human immunoglobulin chl genes and their hypermutation.
A:Reference number: S40312; MUID:94080891
A:Accession: S40320
A:Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 1-96 <KLE>
A:Cross-references: EMBL:X72430; NID:g441328; PIDN:CAAS1098.1; PID:g441329
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 100.0%; Score 33; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 59 FTLKISR 65

RESULT *15

S24501
Ig kappa chain V region - mouse
C:Species: Mus musculus (house mouse)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jan-2000
C:Accession: S24501
R:Kaartinen, M.
submitted to the EMBL Data Library, October 1991

A:Reference number: S24501
A:Accession: S24501
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-99 <KAA>
A:Cross-references: EMBL:X66638
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:16-95/Domain: immunoglobulin homology <IMM>

Query Match 100.0%; Score 33; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
|||||||
Db 76 FTLKISR 82

Search completed: July 15, 2002, 13:01:01
Job time: 467 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:23:32 ; Search time 44.9 Seconds

(without alignments)
6.036 Million cell updates/sec

Title: US-09-712-819A-6

Perfect score: 33

Sequence: 1 FTUKISR 7

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	112	1 KV2C_HUMAN	P01616 homo sapien
2	33	100.0	113	1 KV2D_HUMAN	P01617 homo sapien
3	33	100.0	113	1 KV2G_MOUSE	P01631 mus musculu
4	33	100.0	115	1 KV2A_HUMAN	P01614 homo sapien
5	33	100.0	117	1 KV2E_HUMAN	P06309 homo sapien
6	33	100.0	133	1 KV2F_HUMAN	P06310 homo sapien
7	30	90.9	112	1 KV2D_MOUSE	P01629 mus musculu
8	30	90.9	113	1 KV2B_HUMAN	P01615 homo sapien
9	30	90.9	113	1 KV2E_MOUSE	P03976 mus musculu
10	30	90.9	113	1 KV2F_MOUSE	P01630 mus musculu
11	30	90.9	512	1 ANM3_HUMAN	O60678 homo sapien
12	30	90.9	772	1 C1PB_CLOTT	O01866 clostridium
13	30	90.9	1853	1 C1PA_CLOTT	O06851 clostridium
14	29	87.9	112	1 KV2A_MOUSE	P01626 mus musculu
15	29	87.9	113	1 KV2C_MOUSE	P01627 mus musculu
16	29	87.9	120	1 KV2B_MOUSE	P01627 mus musculu
17	29	87.9	834	1 PLSB_PSEAR	O9hwv7 pseudomonas
18	28	84.8	259	1 Y068_CHITR	O84071 chlamydia t
19	28	84.8	259	1 Y339_CHLMD	O9pxk1 chlamydia m
20	28	84.8	711	1 CAD2_LISMO	O60048 listeria m
21	28	84.8	745	1 KATB_ARATH	P46864 arabidopsis
22	27	81.8	100	1 KV3C_HUMAN	P01619 homo sapien
23	27	81.8	108	1 KV3A_HUMAN	P01620 homo sapien
24	27	81.8	109	1 KV3D_HUMAN	P01622 homo sapien
25	27	81.8	109	1 KV3E_HUMAN	P01623 homo sapien
26	27	81.8	109	1 KV3F_HUMAN	P04206 homo sapien
27	27	81.8	109	1 KV3G_HUMAN	P47105 saccharomyc
28	27	81.8	127	1 YJ07_YEAST	P04207 homo sapien
29	27	81.8	129	1 KV3H_HUMAN	P18135 homo sapien
30	27	81.8	129	1 KV3I_HUMAN	P18136 homo sapien
31	27	81.8	129	1 KV3K_HUMAN	P18136 homo sapien
32	27	81.8	295	1 YH08_YEAST	P38805 saccharomyc
33	27	81.8	299	1 AFNC_BUFMA	P43002 bufo marinu

34	27	81.8	362	1 RECF_PROMI	P22839 proteus mir
35	27	81.8	1231	1 KF4A_MOUSE	P33174 mus musculu
36	27	81.8	1232	1 KF4A_HUMAN	O95239 homo sapien
37	27	81.8	1541	1 NX3A_HUMAN	O9Y4C0 homo sapien
38	27	81.8	1578	1 NX3A_RAT	O07310 rattus norv
39	27	81.8	2068	1 T2D1_DROME	P51123 drosophila
40	26	78.8	108	1 KVL1_CANFA	P01618 canis fami1
41	26	78.8	172	1 R1M1_STRPY	P58183 streptococ
42	26	78.8	194	1 Y4H0_RHISN	P50361 rhizobium s
43	26	78.8	184	1 Y994_MERTH	O27075 methanobact
44	26	78.8	429	1 YFAY_ECOLI	P76470 escherichia
45	26	78.8	670	1 YBY2_SCHPO	Q10199 schizosacch

ALIGNMENTS

```

RESULT 1
KV2C_HUMAN          STANDARD:      PRT:      112 AA.
ID  KV2C_HUMAN
AC  P01616:
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Ig kappa chain V-II region ML.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OK  NCBI_TaxID:9606;
RN  (1)
RA  SEQUENCE.
RP  Dreyer W.J., Gray W.R., Hood L.E.;
RT  "The genetic, molecular, and cellular basis of antibody formation:
RL  some facts and a unifying hypothesis.";
CC  Cold Spring Harb. Symp. Quant. Biol. 32:353-367(1967).
CC  -1- MISCELLANEOUS: THIS C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
DR  PIR: A01887; K2HUML.
DR  InterPro: IPR003006; IG_MHC.
DR  InterPro: IPR003596; IG_V.
DR  Pfam: PF00047; Ig_1.
DR  SMART: SM00406; IGV_1.
KW  Immunoglobulin V region; Bence-Jones protein.
FT  DOMAIN 1 23 FRAMEWORK-1.
FT  DOMAIN 2 38 COMPLEMENTARITY-DETERMINING-1.
FT  DOMAIN 3 53 COMPLEMENTARITY-DETERMINING-2.
FT  DOMAIN 4 60 COMPLEMENTARITY-DETERMINING-3.
FT  DOMAIN 5 92 FRAMEWORK-3.
FT  DOMAIN 6 101 COMPLEMENTARITY-DETERMINING-4.
FT  DOMAIN 7 102 111 FRAMEWORK-4.
FT  DISULFID 23 92 BY SIMILARITY.
FT  NON_TER 112 112.
SQ  SEQUENCE 112 AA: 12055 MW; E5B222EFA7ABE481 CRC64;

Query Match          100.0%; Score 33; DB 1; Length 112;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY  1 FTUKISR 7
DB  75 FTUKISR 81

RESULT 2
KV2D_HUMAN          STANDARD:      PRT:      113 AA.
ID  KV2D_HUMAN
AC  P01617:
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  Ig kappa chain V-II region TEW.
OS  Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (BENCE-JONES PROTEIN TEM).
RX MEDLINE-74148480; PubMed-4596149;
RA Putnam F.W., Whitley E.J. Jr., Paul C., Davidson J.N.;
RT "Amino acid sequence of a kappa Bence Jones protein from a case of
RT primary amyloidosis."
RL Biochemistry 12:3763-3780(1973).
RN [2]
RP SEQUENCE OF 1-27 (AMYLOID PROTEIN TEM).
RX MEDLINE-73166638; PubMed-4700495;
RA Glenner G.G.;
RA Terry W.D., Page D.L., Kimura S., Isobe T., Osseman E.F.,
RT "Structural identity of Bence Jones and amyloid fibril proteins in a
RT patient with plasma cell dyscrasia and amyloidosis."
RL J. Clin. Invest. 52:1276-1281(1973).
CC -1- MISCELLANEOUS: THE MAJOR AMYLOID PROTEIN APPEARS TO BE IDENTICAL
CC WITH THE BENCE JONES PROTEIN ISOLATED FROM THE SAME PATIENT.
CC -1- MISCELLANEOUS: THIS PROTEIN WAS ISOLATED FROM THE URINE OF A
CC PATIENT WITH PLASMA CELL DYSKRASIA AND AMYLOIDOSIS.
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (1,2)
CC MARKER.
DR PIR: A01888; K2HUTM.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KM Immunoglobulin V region; Bence-Jones protein; Amyloid.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 39 FRAMEWORK-2.
FT DOMAIN 3 54 FRAMEWORK-3.
FT DOMAIN 4 61 FRAMEWORK-4.
FT DOMAIN 5 62 FRAMEWORK-5.
FT DOMAIN 6 62 FRAMEWORK-6.
FT DOMAIN 7 94 FRAMEWORK-7.
FT DOMAIN 8 102 FRAMEWORK-8.
FT DOMAIN 9 103 FRAMEWORK-9.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 113
SQ SEQUENCE 113 AA; 12316 MW; 0C3C38F81F1843CA CRC64;

Query Match 100.0%; Score 33; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
DB 76 FTLKISR 82

RESULT 3
KV2G_MOUSE STANDARD; PRT; 113 AA.
ID KV2G_MOUSE
AC P01631.
DR 21-JUL-1986 (Rel. 01, Created)
DR 21-JUL-1986 (Rel. 01, Last sequence update)
DR 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 26-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RC STRAIN=A/J.
RX MEDLINE-83178921; PubMed-6404298;
RA Novotny J., Margolis M.N.;
RT "Amino acid sequence of the light chain variable region from a mouse
RT anti-digoxin hybridoma antibody."
RL Biochemistry 22:1153-1158(1983).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN ICG2A HYBRIDOMA
CC PROTEIN THAT BINDS DIGOXIN.

```

```

DR PIR: A01914; KWS26.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KM Immunoglobulin V region; Monoclonal antibody; Hybridoma.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 2 39 FRAMEWORK-2.
FT DOMAIN 3 54 FRAMEWORK-3.
FT DOMAIN 4 61 FRAMEWORK-4.
FT DOMAIN 5 62 FRAMEWORK-5.
FT DOMAIN 6 62 FRAMEWORK-6.
FT DOMAIN 7 94 FRAMEWORK-7.
FT DOMAIN 8 102 FRAMEWORK-8.
FT DOMAIN 9 103 FRAMEWORK-9.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 113
SQ SEQUENCE 113 AA; 12273 MW; F9F39CE949A84C2A CRC64;

```

```

Query Match 100.0%; Score 33; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FTLKISR 7
DB 76 FTLKISR 82

```

```

RESULT 4
KV2A_HUMAN STANDARD; PRT; 115 AA.
ID KV2A_HUMAN
AC P01614.
DR 21-JUL-1986 (Rel. 01, Created)
DR 21-JUL-1986 (Rel. 01, Last sequence update)
DR 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region Cum.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE-68242259; PubMed-5586923;
RA Hilschmann N.;
RT "The complete amino acid sequence of Bence Jones protein Cum (kappa-
RT type)."
RL Hoppe-Seyler's Z. Physiol. Chem. 348:1718-1722(1967).
RN [2]
RP REVISIONS TO 50; 52; 96 AND 97.
RX MEDLINE-70063440; PubMed-4188189;
RA Hilschmann N.;
RT "Molecular basis of antibody formation."
RL Naturwissenschaften 56:195-205(1969).
CC -1- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -1- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR: A01885; K2HUCM.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KM Immunoglobulin V region; Bence-Jones protein.
FT DISULFID 24 95 BY SIMILARITY.
FT NON_TER 115
SQ SEQUENCE 115 AA; 12676 MW; 59E9P90A379569EC CRC64;

```

```

Query Match 100.0%; Score 33; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
DB 78 FTLKISR 84

```

```

RESULT 5
KV2E_HUMAN STANDARD: PRT; 117 AA.
AC P06509;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region GM607 precursor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84191506; PubMed=6325927;
RA Klobeck H.G., Solomon A., Zachau H.G.;
RT "Contribution of human V kappa II germ-line genes to light-chain
RT diversity";
RL Nature 309:73-76(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 200009; -; NOT_ANNOTATED_CDS.
DR PIR; A01889; K2HUGM.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region; Signal.
KV2E_NON_TER 1 1
KV2E_SIGNAL 1 4
FT CHAIN 5 117 IG KAPPA CHAIN V-II REGION GM607.
FT DOMAIN 5 27 FRAMEWORK-1.
FT DOMAIN 28 43 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 44 58 FRAMEWORK-2.
FT DOMAIN 59 65 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 66 97 FRAMEWORK-3.
FT DOMAIN 98 106 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 107 116 FRAMEWORK-4.
FT DISULFID 27 97 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12664 MW; 92C57DC719E558B1 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
DB 80 FTLKISR 86

RESULT 6
KV2E_HUMAN STANDARD: PRT; 133 AA.
AC P06310;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region RPMI 6410 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=86041852; PubMed=2997711;
RA Klobeck H.G., Meindl A., Combratio G., Solomon A., Zachau H.G.;
RT "Human immunoglobulin kappa light chain genes of subgroups II and
RT III";
RL Nucleic Acids Res. 13:6499-6513(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; 200020; CAA77315.1; -.
DR PIR; A01890; K2HURP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KV2E_NON_TER 1 20
KV2E_SIGNAL 1 20
FT CHAIN 21 133 IG KAPPA CHAIN V-II REGION RPMI 6410.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 59 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 60 74 FRAMEWORK-2.
FT DOMAIN 75 81 FRAMEWORK-3.
FT DOMAIN 82 113 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 114 122 FRAMEWORK-4.
FT DOMAIN 123 132 COMPLEMENTARITY-DETERMINING-3.
FT DISULFID 43 113 BY SIMILARITY.
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 14707 MW; 513CCAF36730099E CRC64;

Query Match 100.0%; Score 33; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKISR 7
DB 96 FTLKISR 102

RESULT 7
KV2D_MOUSE STANDARD: PRT; 112 AA.
AC P01629;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region 2S1.3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=83055101; PubMed=7141411;
RA Herbst H., Chang J.Y., Aebersold R., Braun D.G.;
RT "Murine VK25 isotypic sequence: monoclonal antibody 2S1.3 specific for
RT the group A streptococcal polysaccharide.";
RL Hoppe-seyler's Z. Physiol. Chem. 363:1069-1076(1982)
CC -I- ANTICELLULOUS: THIS CHAIN IS FROM A HYBRIDOMA-DERIVED MONOCLONAL
CC ANTIBODY AGAINST THE STREPTOCOCCAL GROUP A POLYSACCHARIDE.
DR PIR; A01911; KVMSS1.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KV2D_NON_TER 1 23
KV2D_DOMAIN 1 23 FRAMEWORK-1.

```

FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 40 54 FRAMEWORK-2.
 FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 62 93 FRAMEWORK-3.
 FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 103 112 FRAMEWORK-4.
 FT DISULFID 23 93 BY SIMILARITY.
 FT NON_TER 112 112
 SQ SEQUENCE 112 AA: 12221 MW: BDSEF5E6D789F8EC CRC64;

Query Match 90.9%; Score 30; DB 1; Length 112;
 Best Local Similarity 85.7%; Pred. No. 1.8;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLKISR 7
 111:111
 Db 76 FTLKISR 82

RESULT 8

KV2B_HUMAN STANDARD; PRT; 113 AA.
 AC P01615;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-II region FR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Plimates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=76253627; PubMed=821524;

RA Riesen W.F., Jaton J.-C.;
 RT "Variable region sequence of the light chain from a Waldenström's
 RT B-lymphoma with specificity for phosphorylcholine."
 RL Biochemistry 15:3829-3833(1976).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
 CC MACROGLOBULIN THAT BINDS PHOSPHORYLCHOLINE.
 DR PIR: A01886; K2HUF.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 KM Immunoglobulin V region.

FT DOMAIN 1 23 FRAMEWORK-1.
 FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 40 54 FRAMEWORK-2.
 FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 62 93 FRAMEWORK-3.
 FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 103 112 FRAMEWORK-4.
 FT DISULFID 23 93 BY SIMILARITY.
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA: 12660 MW: OCDDA39E46DB96BC CRC64;

Query Match 90.9%; Score 30; DB 1; Length 113;
 Best Local Similarity 85.7%; Pred. No. 1.8;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLKISR 7
 111:111
 Db 76 FTLKISR 82

RESULT 9

KV2E_MOUSE STANDARD; PRT; 113 AA.
 AC P03976;
 DT 23-OCT-1986 (Rel. 02, Created)

DT 23-OCT-1986 (Rel. 02, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-II region 17S29.1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.

RC TISSUE=Hydridoma;
 RX MEDLINE=85128968; PubMed=6441768;
 RA Aebersold R., Herbst H., Grutter T., Chang J.Y., Braun D.G.;
 RT "Murine V kappa 25 and V kappa 27 amino-acid sequences of C57Bl/6
 RT origin: monoclonal antibodies 17S29.1 and 22S25.1 specific for the
 RT group A streptococcal polysaccharide."
 RL Hoppe-Seivler's Z. Physiol. Chem. 365:1375-1383(1984).
 CC -1- MISCELLANEOUS: ANTI-STREPTOCOCCAL GROUP A CARBOHYDRATE ANTIBODY.
 CC PIR: A01912; KVM517.

DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.
 KM Immunoglobulin V region; Hydridoma.
 FT DOMAIN 1 23 FRAMEWORK-1.
 FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
 FT DOMAIN 40 54 FRAMEWORK-2.
 FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
 FT DOMAIN 62 93 FRAMEWORK-3.
 FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
 FT DOMAIN 103 112 FRAMEWORK-4.
 FT DISULFID 23 93 BY SIMILARITY.
 FT NON_TER 113 113
 SQ SEQUENCE 113 AA: 12390 MW: 4E93797046F8D33 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 113;
 Best Local Similarity 85.7%; Pred. No. 1.8;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLKISR 7
 111:111
 Db 76 FTLKISR 82

RESULT 10

KV2F_MOUSE STANDARD; PRT; 113 AA.
 ID KV2E_MOUSE
 AC P01630;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-II region 7S34.1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=83256427; PubMed=6409088;
 RA Chang J.-Y., Herbst H., Aebersold R., Braun D.G.;
 RT "A new isotype sequence (V kappa 27) of the variable region of kappa-
 RT light chains from a mouse hydridoma-derived anti-(streptococcal group
 RT A polysaccharide) antibody containing an additional cysteine residue-
 RT Application of the dimethylaminoazobenzene isothiocyanate technique
 RT for the isolation of peptides."
 RL Biochem. J. 211:173-180(1983).

CC -1- MISCELLANEOUS: THIS CHAIN IS FROM A HYBRIDOMA-DERIVED MONOCLONAL
 CC ANTIBODY AGAINST THE STREPTOCOCCAL GROUP A POLYSACCHARIDE.
 CC PIR: A01913; KVM57S.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_V.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00406; IGV; 1.

Immunoglobulin V region; Monoclonal antibody; Hybridoma.

FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 40 54 FRAMEWORK-2.
FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 62 93 FRAMEWORK-3.
FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 103 112 FRAMEWORK-4.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12496 MW; 42C019D10ADA3C91 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 113;
Best Local Similarity 85.7%; Pred. No. 1.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTUKISR 7
1111111
Db 76 FTUKISR 82

RESULT 11
ANN3 HUMAN STANDARD; PRT; 512 AA.
AC 060678;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein arginine N-methyltransferase 3 (EC 2.1.1.-) (Fragment).
GN PRMT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-98307932; PubMed-9642256;
RA Tang J., Gary J.D., Clarke S., Herschman H.R.;
RT PRMT 3, a type I protein arginine N-methyltransferase that differs
RT from PRMT1 in its oligomerization, subcellular localization,
RT substrate specificity, and regulation.";
RL J. Biol. Chem. 273:16935-16945(1998).
CC -1- FUNCTION: METHYLATES (MONO AND ASYMMETRIC DIMETHYLATION) THE
CC GUANIDINO NITROGENS OF ARGINYL RESIDUES IN SOME PROTEINS.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PROTEIN ARGININE N-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF059531; AAC39837.1; -.
DR MIM: 603190; -.
DR InterPro: IPR000051; SAM_bind.
DR InterPro: IPR000822; ZnF-C2H2.
DR SMART: SM00355; ZnF_C2H2_1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW transferase; Methyltransferase; Zinc-finger.
FT NON_TER 1 1
FT ZN_FING 29 52 C2H2-TYPE.
SQ SEQUENCE 512 AA; 58098 MW; 260DC4EB25162a18 CRC64;

Query Match 90.9%; Score 30; DB 1; Length 512;
Best Local Similarity 85.7%; Pred. No. 8.8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTUKISR 7
1111111
Db 419 FTUKISR 425

RESULT 12
CIPB_CLOTM STANDARD; PRT; 772 AA.
ID CIPB_CLOTM
AC 001866;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cellulosomal scaffolding protein B (Cellulosomal glycoprotein S1/S2)
DE (Cellulose integrating protein B) (Fragment).
GN CIPB.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y5;
RX MEDLINE-93146373; PubMed-1490597;
RA Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
RA Gilbert H.J.;
RT Identification of the cellulose-binding domain of the cellulosome
RT subunit S1 from Clostridium thermocellum Y5.";
RL FEMS Microbiol. Lett. 78:181-186(1992).
CC -1- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSE. IT
CC PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
CC CELLULOYTIC ENZYMES PROBABLY THROUGH THE BINDING OF THE NINE
CC REPEATED DOMAINS WITH THE DUPLICATED SEGMENT OF 24 AMINO ACIDS
CC PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSE.
CC -1- SUBCELLULAR LOCATION: CELL SURFACE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X68233; CAA48312.1; -.
DR HSP: 006851; INRC.
DR InterPro: IPR001956; CBD_3.
DR InterPro: IPR002102; Cohesin.
DR InterPro: IPR002105; Dockerin_1.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00942; CBD_3; 1.
DR Pfam: PF00963; Cohesin; 3.
DR Pfam: PF00404; Dockerin_1; 2.
DR ProDom: PD001947; CBD_3; 1.
DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE: PS00448; CLOS_CELLULOSE_RPT; 2.
KW Cellulose degradation; Cell wall; Glycoprotein; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 607 3 x 147 AA APPROXIMATE REPEATS.
FT REPEAT <1 80 1 (PARTIAL).
FT REPEAT 94 240 2.
FT REPEAT 462 607 3.
FT DOMAIN 81 93 PRO/THR-RICH.
FT DOMAIN 241 272 PRO/THR-RICH.
FT DOMAIN 273 439 CELLULOSE-BINDING.
FT DOMAIN 440 761 PRO/THR-RICH.
FT DOMAIN 710 733 2 x 24 AA APPROXIMATE REPEATS.
FT REPEAT 743 766 2-1.
SQ SEQUENCE 772 AA; 82491 MW; BBF06DE5E094FE10 CRC64;

Query Match 90.9% Score 30; DB 1; Length 772;
 Best Local Similarity 85.7%; Pred. No. 13;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
 Db 672 YTLKISR 678

RESULT 13
 CIPA_CLOTM STANDARD; PRT: 1853 AA.
 AC 006851;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Cellulosomal scaffolding protein A precursor (cellulosomal
 glycoprotein S1/S5) (Cellulose Integrating protein A) (Cohesin).
 GN CIPA.
 OS Clostridium thermocellum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 CC Clostridium.
 OX NCBI_Taxid-1515;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43.
 RC STRAIN-ATCC 27405;
 RX MEDLINE-93302508; PubMed-8316083;
 RA Gerguross U.T., Romaniec M.P.M., Kobayashi T., Huaksson N.S.,
 RA Demain A.L.;
 RT "Sequencing of a Clostridium thermocellum gene (cipa) encoding the
 RT cellulosomal SL-protein reveals an unusual degree of internal
 RT homology."
 RL Mol. Microbiol. 8:325-334(1993).
 RN [2]
 RP SEQUENCE OF 1820-1853 FROM N.A.
 RX MEDLINE-93209931; PubMed-8458832;
 RA Fujino T., Beguin P., Aubert J.-P.;
 RT "Organization of a Clostridium thermocellum gene cluster encoding the
 RT cellulosomal scaffolding protein Cipa and a protein possibly involved
 RT in attachment of the cellulosome to the cell surface."
 RL J. Bacteriol. 175:1891-1899(1993).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
 RX MEDLINE-97238934; PubMed-9083107;
 RA Shimon L.J.W., Bayer E.A., Moray E., Lamed R., Yaron S., Shoham Y.,
 RA Frolov F.;
 RT "A cohesin domain from Clostridium thermocellum: the crystal
 RT structure provides new insights into cellulosome assembly."
 RL Structure 5:381-390(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
 RX MEDLINE-97076134; PubMed-8918451;
 RA Tormo J., Lamed R., Chirino A.J., Moray E., Bayer E.A., Shoham Y.,
 RA Steitz T.A.;
 RT "Crystal structure of a bacterial family-III cellulose-binding
 RT domain: a general mechanism for attachment to cellulose."
 RL EMBO J. 15:5739-5751(1996).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361.
 RX MEDLINE-98022914; PubMed-9402065;
 RA Tavares G.A., Beguin P., Alzari P.M.;
 RT "The crystal structure of a type I cohesin domain at 1.7-A
 RT resolution."
 RL J. Mol. Biol. 273:701-713(1997).
 CC -1- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSOME. IT
 CC PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
 CC CELLULOLYTIC ENZYMES.
 CC -1- SUBCELLULAR LOCATION: CELL SURFACE.
 CC -1- DOMAIN: CONTAINS 9 COPIES OF A DOMAIN (COHESIN) OF ABOUT 150
 CC RESIDUES THAT BINDS TO THE DOCKERIN DOMAIN BORN BY THE CATALYTIC
 CC COMPONENTS OF THE CELLULOSOME.
 CC -1- DOMAIN: A 24 RESIDUES DOMAIN IS REPEATED TWICE IN THIS ENZYME AS
 CC WELL AS IN OTHER C.THERMOCELLUM CELLULOSOME ENZYMES. THIS DOMAIN

MAY FUNCTION AS THE BINDING LIGAND FOR THE SL COMPONENT.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib.ch).

 CC EMBL: L08665; NOT_ANNOTATED_CDS.
 CC EMBL: X67506; CAA47840.1; -.
 CC PIR: S36859; S36859.
 CC PDB: 1ANU; 23-JUL-97.
 CC PDB: 1AOH; 08-JUL-98.
 CC PDB: 1NBC; 26-SEP-97.
 CC InterPro: IPR001956; CBD_3.
 CC InterPro: IPR002102; Cohesin.
 CC InterPro: IPR002105; Cohesin.
 CC InterPro: IPR002048; EF-hand.
 CC Pfam: PF00942; CBD_3; 1.
 CC Pfam: PF00963; Cohesin; 9.
 CC Pfam: PF00404; Dockerin_1; 2.
 CC ProDom: PD001947; CBD_3; 1.
 CC ProSite: PS00018; EF_HAND; UNKNOWN_1.
 CC ProSite: PS00448; CLOS_CELLULOSOME_RPT; 2.
 CC Cellulose degradation; Cell wall; Glycoprotein; Repeat; Signal;
 KW 3D-structure.
 FT SIGNAL 1 28
 FT CHAIN 29 1853 CELLULOSOMAL SCAFFOLDING PROTEIN A.
 FT DOMAIN 364 522 CELLULOSE-BINDING (BY SIMILARITY).
 FT REPEAT 54 1694 9 X 150 AA APPROXIMATE TANDEM REPEATS.
 FT REPEAT 29 182 1-1.
 FT REPEAT 183 322 1-2.
 FT REPEAT 560 704 1-3.
 FT REPEAT 724 866 1-4.
 FT REPEAT 889 1031 1-5.
 FT REPEAT 1054 1196 1-6.
 FT REPEAT 1219 1361 1-7.
 FT REPEAT 1384 1526 1-8.
 FT REPEAT 1548 1690 1-9.
 FT DOMAIN 1791 1846 2 X 24 AA APPROXIMATE REPEATS.
 FT REPEAT 1791 1814 2-1.
 FT REPEAT 1824 1847 2-2.
 FT CONFLICT 1615 1615 A -> AA (IN REF. 1).
 SO SEQUENCE 1853 AA; 196831 MW; 3ABDDC03ABFC5372 CRC64;

Query Match 90.9% Score 30; DB 1; Length 1853;
 Best Local Similarity 85.7%; Pred. No. 33;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
 Db 1753 YTLKISR 1759

RESULT 14
 KV2A_MOUSE STANDARD; PRT: 112 AA.
 AC P01626;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig kappa chain V-II region MOPC 167.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 OX NCBI_Taxid-10090;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-79000273; PubMed-99160;
 RA Rudnikoff S., Potter M.;

RT "Kappa Chain variable region from M167, a phosphorylcholine binding
RT myeloma protein."
RL Biochemistry 17:2703-2707(1978).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE. THE SEQUENCE OF THE V REGION OF THE HEAVY
CC CHAIN HAS ALSO BEEN DETERMINED.
DR PIR: A01908; KWS16.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00406; IGv; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 40 54 FRAMEWORK-2.
FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 62 93 FRAMEWORK-3.
FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 103 112 FRAMEWORK-4.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 112 112
SQ SEQUENCE 112 AA; 12349 MW; A58EDFD6404B9726 CRC64;

Query Match 87.9%; Score 29; DB 1; Length 112;
Best Local Similarity 85.7%; Pred. No. 3.2;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 FTLKISR 7
|||:||||
Db 76 FTLEISR 82

RESULT 15
KVZC_MOUSE
ID KVZC_MOUSE STANDARD; PRT; 113 AA.
AC P01628;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-II region MOPC 511.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=81052016; PubMed=6776396;
RA Appella E.;
RT "Amino acid sequence of the light chain variable region of M511, a
RT phosphorylcholine-binding murine myeloma protein."
RL Mol. Immunol. 17:711-718(1980).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS PHOSPHORYLCHOLINE.
CC PIR: A01910; KWS51.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003596; IG_V.
DR Pfam: PF00047; IG; 1.
DR SMART: SM00406; IGv; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23 FRAMEWORK-1.
FT DOMAIN 24 39 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 40 54 FRAMEWORK-2.
FT DOMAIN 55 61 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 62 93 FRAMEWORK-3.
FT DOMAIN 94 102 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 103 112 FRAMEWORK-4.
FT DISULFID 23 93 BY SIMILARITY.
FT NON_TER 113 113
SQ SEQUENCE 113 AA; 12496 MW; EFB0DC4DA2BD3450 CRC64;

Query Match 87.9%; Score 29; DB 1; Length 113;

Best Local Similarity 85.7%; Pred. No. 3.3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 1 FTLKISR 7
|||:||||
Db 76 FTLEISR 82

Search completed: July 15, 2002, 13:23:33
Job time: 1449 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 15, 2002, 13:22:45 ; Search time 172.49 Seconds
(without alignments)
7.020 Million cell updates/sec

Title: US-09-712-819a-6

Perfect score: 33

Sequence: 1 FTLKISR 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	104	11 Q9JL82	Q9JL82 mus musculus
2	33	100.0	114	4 Q9UL80	Q9UL80 homo sapien
3	33	100.0	238	11 Q9JL87	Q9JL87 mus musculus
4	30	90.9	471	5 Q9XU13	Q9XU13 caenorhabdi
5	30	90.9	638	2 Q46453	Q46453 clostridium
6	30	90.9	3848	5 Q76737	Q76737 dictyosteli
7	28	84.8	31	2 Q9R4X1	Q9R4X1 treponema d
8	28	84.8	80	2 Q9FDR9	Q9FDR9 corynebacte
9	28	84.8	80	2 Q9FPA7	Q9FPA7 brevibacter
10	28	84.8	118	16 Q9JL82	Q9JL82 mus musculus
11	28	84.8	285	16 Q9RW56	Q9RW56 streptococc
12	28	84.8	310	11 Q9CU48	Q9CU48 mus musculus
13	28	84.8	327	5 Q9S735	Q9S735 caenorhabdi
14	28	84.8	346	11 Q9CU16	Q9CU16 mus musculus
15	28	84.8	347	10 Q93XK8	Q93XK8 zea mays (m
16	28	84.8	389	16 Q25171	Q25171 helicobacte

17	28	84.8	389	16 Q9ZK68	Q9ZK68 helicobacte
18	28	84.8	399	2 Q56257	Q56257 treponema d
19	28	84.8	449	5 Q9VC17	Q9VC17 drosophila
20	28	84.8	477	5 Q9T677	Q9T677 caenorhabdi
21	28	84.8	546	4 Q9NM40	Q9NM40 homo sapien
22	28	84.8	564	16 Q9JYP8	Q9JYP8 streptococc
23	28	84.8	733	16 Q9Z6C7	Q9Z6C7 rhizobium m
24	28	84.8	744	10 Q9T047	Q9T047 arabidopsis
25	28	84.8	943	2 Q52149	Q52149 synecchococ
26	28	84.8	980	10 Q9JL82	Q9JL82 mus musculus
27	28	84.8	1015	10 Q9JL82	Q9JL82 mus musculus
28	28	84.8	1490	5 Q9T8X6	Q9T8X6 drosophila
29	28	81.8	51	17 Q973A6	Q973A6 sulfolobus
30	27	81.8	109	4 Q9UL86	Q9UL86 homo sapien
31	27	81.8	109	4 Q9UL78	Q9UL78 homo sapien
32	27	81.8	147	3 Q08915	Q08915 saccharomye
33	27	81.8	171	10 Q9C8P6	Q9C8P6 arabidopsis
34	27	81.8	204	16 Q9KM12	Q9KM12 vldrio chol
35	27	81.8	219	16 Q9RT85	Q9RT85 deinococcus
36	27	81.8	294	16 P72930	P72930 synecchocyst
37	27	81.8	315	10 Q9ZU65	Q9ZU65 arabidopsis
38	27	81.8	316	2 Q9AD85	Q9AD85 ethiopia c
39	27	81.8	367	16 Q986H5	Q986H5 rhizobium l
40	27	81.8	389	17 Q57712	Q57712 pyrococcus
41	27	81.8	437	16 Q9CUX3	Q9CUX3 pasteurella
42	27	81.8	443	2 Q32628	Q32628 haemophilus
43	27	81.8	443	2 Q32624	Q32624 haemophilus
44	27	81.8	488	12 Q67635	Q67635 gallid herp
45	27	81.8	540	17 Q971R4	Q971R4 sulfolobus

ALIGNMENTS

RESULT	ID	Q9JL82	PRELIMINARY:	PRT:	104 AA.
AC	Q9JL82:				
DT	01-OCT-2000	(TREMBL:15, Created)			
DT	01-OCT-2000	(TREMBL:15, Last sequence update)			
DT	01-DEC-2001	(TREMBL:19, Last annotation update)			
DE	ANTI-MYOSIN IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION				
DE	(FRAGMENT).				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_Taxid=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=BAIB/C;				
RX	MEDLINE=20448942; PubMed=10992488;				
RA	Maiktel S., Liao L., Cunningham M.W., Diamond B.;				
RT	"T-Cell-dependent antibody response to the dominant epitope of				
RT	streptococcal polysaccharide, N-acetyl-glucosamine, is cross-reactive				
RT	with cardiac myosin."				
RL	Infect. Immun. 68:5803-5808(2000).				
DR	EMBL; AF260024; AAF69322.1; .				
DR	HSSP; P01607; IREI.				
DR	InterPro; IPR003006; Ig_MHC.				
DR	InterPro; IPR003596; Ig_V.				
DR	Pfam; PF00047; Ig_1.				
DR	SMART; SM00406; IGV_1.				
FT	NON_TER	1			
FT	NON_TER	104			
SQ	SEQUENCE	104 AA; 11360 MW; SDA8BBD5F0AAAE CRC64;			

Query Match 100.0%; Score 33; DB 11; Length 104;
Best local similarity 100.0%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FTLKISR 7
|||||||

Db 68 FTLKISR 74

RESULT 2
Q9UL80 PRELIMINARY: PRT: 114 AA.

AC Q9UL80: 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M., Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal fetus."
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL: AF035034; AAD56270.1; -.
DR HSP: P80362; 1MTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 114
SQ SEQUENCE 114 AA; 12775 MW; 070E31E210D1CB01 CRC64;

Query Match 100.0%; Score 33; DB 4; Length 114;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
Db 76 FTLKISR 82

RESULT 3
Q99M37 PRELIMINARY: PRT: 238 AA.

AC Q99M37: 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 26.3 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR. MAP-TGF ALPHA MODEL. 7 MONTHS OLD, GROSS TISSUE.
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC002035; AAH02035.1; -.
DR HSP: P01679; 2FBJ.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_CL.
DR InterPro: IPR003600; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00409; IG; 2.
DR SMART: SM00407; IGC1; 1.
DR SMART: SM00406; IGV; 1.

DR SMART: SM00410; IG_Like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 238 AA; 26344 MW; FB2B06A0B801330A CRC64;

Query Match 100.0%; Score 33; DB 11; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
Db 95 FTLKISR 101

RESULT 4
Q9XUI3 PRELIMINARY: PRT: 471 AA.

AC Q9XUI3: 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ZK39.1 PROTEIN.
GN ZK39.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Kershaw J.K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology."
RT Science 282:2012-2018(1998).
DR EMBL: Z82093; CAB05017.1; -.
SQ SEQUENCE 471 AA; 53577 MW; CB6759DFE21FF959 CRC64;

Query Match 90.9%; Score 30; DB 5; Length 471;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
Db 460 FTLKISR 466

RESULT 5
Q46453 PRELIMINARY: PRT: 638 AA.

AC Q46453: 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CIPA PROTEIN (FRAGMENT).
GN CIPA.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujino T., Beguin P., Aubert J.P.;
RT "Cloning of a Clostridium thermocellum DNA fragment encoding polypeptides that bind the catalytic components of the cellulosome."
RL FEMS Microbiol. Lett. 94:165-170(1992).
RN [2]
RP SEQUENCE FROM N.A.

RX MEDLINE=92394486; PubMed=1521765;
 RA Fujino T., Beguin P., Aubert J.P.;
 RT "Cloning of a Clostridium thermocellum DNA fragment encoding
 RT polypeptide that bind the catalytic components of the cellulosome.";
 RL FEMS Microbiol. Lett. 73:165-170(1992).
 DR EMBL: X67406; CAA47806.1; -.
 DR HSSP: Q06851; IAOH.
 DR InterPro: IPR002102; Cohesin.
 DR InterPro: IPR002105; Dockerin_1.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00963; Cohesin; 3.
 DR Pfam: PF00404; Dockerin_1; 2.
 DR PROSITE: PS00448; CLOS_CELLOSOME_RPT; UNKNOWN_2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 FT NON TER 1 1
 SQ SEQUENCE 638 AA; 68255 MW; 355C17E50AF0CB66 CRC64;

Query Match 90.9%; Score 30; DB 2; Length 638;
 Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
 DB 538 YTLKISR 544

RESULT 6
 ID 076737 PRELIMINARY; PRT; 3848 AA.
 AC 076737;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE TIPC.
 GN TIPC.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX4;
 RA MEDLINE=9331084; PubMed=10402673;
 RA Stege J.T., Laub M.T., Loomis W.F.;
 RT "tip genes act in parallel pathways of early Dictyostelium
 RT development.";
 RL Dev. Genet. 23:64-77(1999).
 DR EMBL: AF079445; AAC31916.1; -.
 SQ SEQUENCE 3848 AA; 440033 MW; 65B671566F7AE534 CRC64;

Query Match 90.9%; Score 30; DB 5; Length 3848;
 Best Local Similarity 85.7%; Pred. No. 3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 7
 DB 886 FTLKISR 892

RESULT 7
 ID 09R4X1 PRELIMINARY; PRT; 31 AA.
 AC 09R4X1;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HEMOLYSIN (FRAGMENT).
 OS Treponema denticola.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_TaxID=158;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=94376628; PubMed=8090078;
 RA Chu L., Holt S.C.;
 RT "Purification and characterization of a 45 kDa hemolysin from
 RT Treponema denticola ATCC 35404.";
 RL Microb. Pathog. 16:197-212(1994).
 SQ SEQUENCE 31 AA; 3650 MW; 054650358CBF24C4 CRC64;

Query Match 84.8%; Score 28; DB 2; Length 31;
 Best Local Similarity 85.7%; Pred. No. 8;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 FTLKISR 7
 DB 5 FTLKISR 11

RESULT 8
 ID 09FDR9 PRELIMINARY; PRT; 80 AA.
 AC 09FDR9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE ATP1 PROTEIN.
 GN ATP1.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC13060;
 RA Sekine H., Yokota A., Tomita F.;
 RT "Nucleotide sequence of atp operon of Corynebacterium glutamicum.";
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB046112; BAB08150.1; -.
 SQ SEQUENCE 80 AA; 9308 MW; 23CC0E33367F8D8 CRC64;

Query Match 84.8%; Score 28; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FTLKISR 6
 DB 36 FTLKISR 41

RESULT 9
 ID 09FAA7 PRELIMINARY; PRT; 80 AA.
 AC 09FAA7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE ATP1 PROTEIN.
 GN ATP1.
 OS Brevibacterium flavum.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriidae;
 OC Actinomycetales; Corynebacterineae.
 OX NCBI_TaxID=92706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC14067;
 RA Sekine H., Tomita F., Yokota A.;
 RT "Nucleotide sequence of atp operon of Brevibacterium flavum.";
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB048368; BAB13353.1; -.
 SQ SEQUENCE 80 AA; 9339 MW; 72CD0E33367EAB33 CRC64;

Query Match 84.8%; Score 28; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKIS 6
 |||||
 Db 36 FTLKIS 41

RESULT 10
 Q99V56 PRELIMINARY; PRT; 118 AA.
 AC Q99V56; ` PRELIMINARY; PRT; 118 AA.
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 GN SA0890 OR SAV1037.
 OS Staphylococcus aureus (strain N315), and
 OC Staphylococcus aureus (strain Mu50).
 OC Bacteria; Firmicutes; Bacilli; Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=158879, 158878;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Ikeno J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ul Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hiraoka H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanetsuna M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL: AP003132; BAB42135.1; -;
 DR EMBL: AP003131; BAB57199.1; -;
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 118 AA; 13229 MW; A895E0245E3BC710 CRC64;

Query Match 84.8%; Score 28; DB 16; Length 118;
 Best Local Similarity 71.4%; Pred. No. 30;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKIS 7
 |||||
 Db 80 FTLKIS 86

RESULT 11
 Q9RWY8 PRELIMINARY; PRT; 285 AA.
 AC Q9RWY8; PRELIMINARY; PRT; 285 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE HYPOTHETICAL 31.2 KDA PROTEIN.
 GN DR0527.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R1;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001911; AAF10109.1; -;
 DR TIGR: DR0527; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 285 AA; 31219 MW; 02FDFAD6D09351E3 CRC64;

Query Match 84.8%; Score 28; DB 16; Length 285;
 Best Local Similarity 71.4%; Pred. No. 73;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FTLKIS 7
 |||||
 Db 25 FTLKIS 31

RESULT 12
 Q9CU48 PRELIMINARY; PRT; 310 AA.
 AC Q9CU48; PRELIMINARY; PRT; 310 AA.
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DE 1110021E09R1K PROTEIN (FRAGMENT).
 GN 1110021E09R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=MEDULLA OBLONGATA;
 RX MEDLINE=21083660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shiba Y., Storch K.-F.,
 RA Suzuki H., Toyono-Oka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RT Nature 409:685-690(2001).
 DR EMBL: AK018126; BAB31082.1; -;
 DR MGD: MGI:1914171; 1110021E09R1K.
 FT NON_TER 310
 SQ SEQUENCE 310 AA; 35234 MW; 017744FA3AD1EBEC CRC64;

Query Match 84.8%; Score 28; DB 11; Length 310;
 Best Local Similarity 71.4%; Pred. No. 79;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 FTLKIS 7
 |||||
 Db 137 FTLKIS 143

RESULT 13